JOURNAL OF MULTIVARIATE ANALYSIS 11, 199–229 (1981)

Stopping Rules and Tactics for Processes Indexed by a Directed Set*

ULRICH KRENGEL

University of Goettingen

AND

LOUIS SUCHESTON

Ohio State University

Communicated by A. Dvoretzky

> It takes little talent to see clearly what lies under one's nose, a good deal of it to know in what direction to point that organ.—W. H. Auden, "The Dyer's Hand," Part I, "Writing," 1962.

Contents. Introduction. 1. Stopping rules and tactics: basic notions. 2. Conditions for representing stopping rules as tactics. 3. Linear embedding of tactics. 4. Applications of the linear embedding theorem. 5. Construction of ordered stopping rules for independent processes. 6. Dominated estimates in terms of stopping rules.

A new notion of tactic for processes indexed by a directed set is introduced. The main theorem, giving conditions under which tactics can be mapped on stopping times on the line, is applied to reduce some optimal stopping problems in the plane to the same problems on the line. In the case of independent random variables, one achieves a nearly complete reduction of the optimal reward problem to the linear case.

Received October 9, 1979; revised September 22, 1980.

AMS 1980 Classification numbers: Primary 60G40, 62K05, 62L15, Secondary 60E15.

Key words and phrases: Stopping rule, tactic, independent random variables, averages of independent identically distributed random variables, optimal reward, maximal inequality, linear embedding.

* This research was supported by the National Science Foundation (USA). The main results were announced in a Note to Comptes Rendus Acad. Sc; Paris, t. 290, pp. 192–196, on January 23, 1980.

0047-259X/81/020199–31$02.00/0

Introduction

Our index set Q is countable, partially ordered, and locally finite. We study stopping problems for stochastic processes $(X_t, t \in Q)$; precise definitions are given in Section 1 below. The new basic notion is that of a *tactic*; each tactic determines a stopping rule (= stopping time), but not all stopping rules come from tactics. To understand the main difference between a tactic and a stopping rule, observe the following difference between the line (our shorthand here for the total order of $\mathbb{N}$) and the plane $(= \mathbb{N} \times \mathbb{N})$, or, more generally, a partially ordered set Q. On the line, the choice at a point t is between stopping and hence accepting the reward X_t, and non-stopping and observing the next random variable X_{t+1}. In general, if one chooses not to stop, then it is still necessary to know which of the immediate successors of X_t one will observe: a tactic gives this information. In the particular case of a "tree," tactics can be for all practical purposes identified with properly measurable single random variables, called *control variables* (Haggstrom [14]). The main results of the present paper require, however, a more delicate definition: a tactic is a set of porperly measurable partitions of the sample space. The notion of a tactic seems important in itself, but the theory given here also sheds now light on stopping rules properly speaking. We show that stopping rules are often given by tactics, and in many cases of interest, tactics can be mapped on the line (the linear embedding theorem), and studied by familiar methods. This enables us to solve, via tactics, several problems concerning stopping rules in two dimensions.

Section 1 gives basic definitions, and also shows that if Q is finite, then an optimal tactic can be obtained by a version of the familiar method called "backward induction." In Section 2 we define a weak independence condition called "conditional qualitative independence" (CQI), and show that under CQI every stopping rule is given by a tactic. In the plane, CQI usually holds, but in three dimensions there are stopping rules not given by tactics even if all the σ-algebras are independent. In Section 3 we prove the linear embedding theorem: for a wide class of processes, in particular for averages of independent identically distributed random variables, the study of the expected reward of a tactic reduces to a linear setting. This is a general device for extending results known for $Q = \mathbb{N}$ to a more general index set. In Section 4 we give some applications: thus the deep theorem of Dvoretzky and Davis asserting the existence of an optimal stopping rule for averages of independent identically distributed random variables extends to $Q = \mathbb{N} \times \mathbb{N}$; so do Wald's equation and Wald's identity [23]. Another application is a positive answer in the case $\mathbb{N} \times \mathbb{N}$ to a question of Cairoli and Gabriel [5] about the degree of integrability needed to insure that the value (the supremum over all stopping rules τ of $E(X_\tau)$) be finite.

In Section 5 we study the case of independent random variables. We achieve here a nearly complete reduction of the optimal reward problem to the linear case: it is shown that if $E(\sup_{t \in Q} X_t^-) < \infty$, then given any stopping rule τ, there exists an *ordered* stopping rule (the range is totally ordered) τ', such that $E(X_\tau) \leqslant E(X_{\tau'})$. An ordered rule can be of course mapped on the line. An application of this is given in the last, the sixth, section, which studies extentions to directed sets of our dominated linear estimates in terms of stopping times (Krengel and Sucheston [15]).

1. Stopping Rules and Tactics: Basic Notions

Q is a countable set with a partial order $\leqslant$. (Q is not assumed filtering to the right.) Elements of Q are denoted $q, r, s, t, u, \ldots$. The smallest element of Q, if it exists, is denoted by p. $\bar{Q} = Q \cup \{\infty\}$, where $\infty \geqslant q$ for all $q \in Q$. $K(t)$ is the set of elements $\leqslant t$; $L(t)$ the set of elements $\geqslant t$. Q is assumed to be *locally finite*, i.e., such that $|K(t)| < \infty$ for each $t \in Q$; $|K(t)|$ is the cardinality of $K(t)$. D_t is the set of elements directly above t, i.e., the set of $u \in Q$ such that $u > t$ and $t \leqslant t' < u$ implies $t' = t$.

$(\Omega, \mathfrak{F}, P)$ is a probability space, and $(\mathfrak{F}_t, t \in Q)$ a *stochastic basis*: a family of sub-σ-algebras of $\mathfrak{F}$ with $\mathfrak{F}_s \subset \mathfrak{F}_t$ if $s \leqslant t$. A *stopping rule*, or *stopping time*, is a map $\sigma: \Omega \to \bar{Q}$, such that $\{\sigma = t\} \in \mathfrak{F}_t$ for all $t \in Q$. A stopping rule is called *finite* if its range is contained in Q. The set of finite stopping rules is denoted by Σ; the set of all stopping rules is $\bar{\Sigma}$.

We now define a tactic. Assume that there is an element $p \in Q$ with $p \leqslant t$ for all $t \in Q$. A *tactic* $\mathscr{H}$ is a family

$$\mathscr{H} = \{H_{st}, s \in Q, t \in \{s\} \cup D_s\}$$

such that (i) $H_{st} \in \mathfrak{F}_s$ for all s, t; (ii) for each fixed s, $\{H_{st}, t \in \{s\} \cup D_s\}$ is a partition of Ω.

A tactic $\mathscr{H}$ generates a unique stopping rule $\tau = \tau_{\mathscr{H}}$ as follows. Let $\omega \in \Omega$; we distinguish two cases.

(a) There exists a *finite* sequence of elements of Q, $p = t(0) < t(1) < \cdots < t(n) = t(n+1)$ with n and $t(k)$ depending on ω, such that $t(i) \in D_{t(i-1)}$, $i = 1, \ldots, n$; this sequence is determined by the requirement that

$$\omega \in \bigcap_{i=0}^{n} H_{t(i)t(i+1)}.$$

Then $\tau_{\mathscr{H}}(\omega) = t(n)$.

(b) There is an *infinite* sequence of elements of Q, $p = t(0) < t(1) < \cdots$, with $t(k)$ depending on ω, such that $t(i) \in D_{t(i-1)}$, $i = 1, 2, \ldots$, and

$$\omega \in \bigcap_{i=0}^{\infty} H_{t(i)t(i+1)}.$$

Then $\tau_{\mathscr{H}}(\omega) = \infty$.

We say that $\tau_{\mathscr{H}}$ is determined, or given, by the tactic $\mathscr{H}$. Clearly, several tactics may determine the same stopping rule. The set of elements of Σ determined by a tactic is denoted by T (capital τ); the set of elements of $\bar{\Sigma}$ determined by a tactic is denoted by $\bar{T}$.

The stopping rules in $\bar{T}$ have the special property that the event that from s one proceeds to the direct successor of s, t, is determined by the information at time s ($H_{st} \in \mathfrak{F}_s$ if $t \in D_s$).

A *random variable* is a map $X: \Omega \to \mathbb{R}$ such that $X^{-1}B \in \mathfrak{F}$ for each Borel set $B \in \mathbb{R}$. An *adapted process* is a set of random variables indexed by Q, $(X_t, t \in Q)$, such that for each t, X_t is measurable with respect to $\mathfrak{F}_t$. Many relations below involving random variables and sets will be assumed to hold only modulo P-null sets; the words *almost surely* (a.s.) may or may not be omitted.

Backward Induction. Let Q be finite, $p \in Q$. Let

$$\begin{aligned}
Q_1 &= \{q \in Q\colon \text{there is no } r > q,\ r \in Q\}\\
Q_2 &= \{q \in Q\colon \text{all } r > q \text{ are in } Q_1\} \backslash Q_1\\
&\vdots\\
Q_k &= \left\{q \in Q\colon \text{all } r > q \text{ are in } \bigcup_{i=1}^{k-1} Q_i\right\} \Big\backslash \bigcup_{i=1}^{k-1} Q_i.
\end{aligned}$$

Given an adapted process $(X_t, t \in Q)$, we define U_t as follows: If $t \in Q_1$, set $U_t = X_t$. If $t \in Q_k$, $k > 1$,

$$U_t = \max\left(X_t, \sup_{t' > t} E^{\mathfrak{F}_t} U_{t'}\right). \tag{1}$$

Clearly, $(U_t, t \in Q)$ is a supermartingale.

LEMMA 1.1. *We have*

$$U_s = \max\left(X_s, \sup_{t \in D_s} E^{\mathfrak{F}_s} U_t\right). \tag{2}$$

Proof. If $s < t < t'$, then $U_t \geq E^{\mathfrak{F}_t} U_{t'}$, hence $E^{\mathfrak{F}_s} U_t \geq E^{\mathfrak{F}_s}(E^{\mathfrak{F}_t} U_{t'}) = E^{\mathfrak{F}_s} U_{t'}$. This implies (2). ∎

Now define by "backward induction" sets H_{st} as follows: If $s \in Q_1$, $H_{ss} = \Omega$.

If $s \in Q_i$ with $i > 1$, set $H_{ss} = \{\omega: X_s \geqslant U_s\}$. Thus H_{ss} is the set where one "chooses to stop at s." Let $(t_1, t_2, \ldots)$ be an enumeration of D_s in an auxiliary total order $\prec$. Set

$$\begin{aligned} H_{st_1} &= \{\omega: U_s = E^{\mathfrak{F}_s} U_{t_1}\} \backslash H_{ss} \\ H_{st_2} &= \{\omega: U_s = E^{\mathfrak{F}_s} U_{t_2}\} \backslash (H_{ss} \cup H_{st_1}) \\ &\vdots \qquad \vdots \qquad \vdots \qquad \vdots \\ H_{st_j} &= \{\omega: U_s = E^{\mathfrak{F}_s} U_{t_j}\} \Big\backslash \Big(H_{ss} \cup \bigcup_{i<j} H_{st_i}\Big), \end{aligned}$$

etc.

For a fixed s, the tactic $\mathscr{H}^s$ on $L(s)$ is defined by

$$\mathscr{H}^s = \{H_{s't'}, s' \geqslant s, t \in \{s'\} \cup D_{s'}\}.$$

Then

$$\begin{aligned} \nu_s = \tau_{\mathscr{H}^s} &= s \qquad \text{on} \quad H_{ss} \\ &= \nu_{t_i} \qquad \text{on} \quad H_{st_i}, t_i \in D_s. \end{aligned}$$

In the following discussion we will identify $\mathscr{H}^s$ with ν_s. Let T_s be the set of all stopping rules $\geqslant s$ and given by tactics; we show that ν_s is optimal in this set, i.e., $EX_{\nu_s} \geqslant EX_\tau$ for all $\tau \in T_s$.

THEOREM 1.2. *We have for each* $s \in Q$

(i) $E^{\mathfrak{F}_s} X_{\nu_s} = U_s$;

(ii) $U_s = \sup_{\tau \in T_s} E^{\mathfrak{F}_s} X_\tau$;

(iii) $E^{\mathfrak{F}_s} X_{\nu_s} \geqslant E^{\mathfrak{F}_s} X_\tau, \tau \in T_s$.

Therefore ν_s *is optimal in* T_s.

Proof. Statement (iii) follows from (i) and (ii). Integrating (iii) one obtains that ν_s is optimal in T_s.

Proof of (i): The statement is obvious for $s \in Q_1$. Now assume that (i) holds for $s \in \bigcup_{i<j} Q_i$. Let $s \in Q_j$. If $A \in \mathfrak{F}_s$ then

$$\int_A X_{\nu_s} = \int_{A \cap H_{ss}} X_{\nu_s} + \sum_{t \in D_s} \int_{A \cap H_{st}} X_{\nu_t}.$$

By the induction hypothesis,

$$\int_{A\cap H_{st}} X_{v_t} = \int_{A\cap H_{st}} U_t = \int_{A\cap H_{st}} E^{\mathfrak{F}_s} U_t.$$

By definition one has $U_s = E^{\mathfrak{F}_s} U_t$ on H_{st}, $t \in D_s$, and $U_s = X_s$ on H_{ss}. Therefore

$$\int_A X_{v_s} = \int_{A\cap H_{ss}} U_s + \sum_{t\in D_s} \int_{A\cap H_{ss}} U_s = \int_A U_s.$$

Proof of (ii): Again assume that the relation holds for all $s \in \bigcup_{i<j} Q_i$, and let $s \in Q_j$. Let

$$V_s = \sup_{\tau \in T_s} E^{\mathfrak{F}_s} X_\tau. \tag{3}$$

Since $v_s \in T_s$, (i) implies that $U_s \leqslant V_s$. Suppose that the inverse inequality fails; then there exists a tactic $\mathscr{H}'$ with $\tau = \tau_{\mathscr{H}'} \in T_s$ such that $E^{\mathfrak{F}_s} X_\tau > U_s$ on a non-null set $A \in \mathfrak{F}_s$. Also, $A \subset \{\tau > s\}$, because $X_s \leqslant U_s$, and $X_\tau = X_s$ on $\{\tau = s\}$ implies $E^{\mathfrak{F}_s} X_\tau = X_s$ on $\{\tau = s\}$.

Then $A \subset \bigcup_{t\in D_s} H'_{st}$ implies that among the H'_{st}, $t \in D_s$, there is at least one, say, H'_{st_j}, such that $P(A \cap H'_{st_j}) > 0$. Let $A' = A \cap H_{st_j}$. Define a new tactic $\mathscr{H}^*$ to agree with $\mathscr{H}$ on A', and such that on $(A')^c$, $\mathscr{H}^*$ stops on t_j, i.e., $\mathscr{H}^*_{t_j t_j} = (A')^c$. If $\tau^* = \tau_{\mathscr{H}^*}$, then

$$\begin{aligned} \tau^* &= \tau \qquad \text{on} \quad A' \\ &= t_j \qquad \text{on} \quad (A')^c. \end{aligned}$$

Since $\tau^* = \tau$ except on the $\mathfrak{F}_s$-measurable set $(A')^c$, $E^{\mathfrak{F}_s} X_\tau > U_s$ on A' implies $E^{\mathfrak{F}_s} X_{\tau^*} > U_s$ on A'. But $\tau^* \in T_{t_j}$, hence by the induction hypothesis $U_{t_j} \geqslant E^{\mathfrak{F}_{t_j}} X_{\tau^*}$. Applying $E^{\mathfrak{F}_s}$ to both sides, we obtain by the supermartingale property of (U_t) that $U_s \geqslant E^{\mathfrak{F}_s} X_{\tau^*}$, a contradiction. ∎

In the case when Q is infinite, we can *define* U_s by the formula

$$U_s = \sup_{\tau \in T_s} E^{\mathfrak{F}_s} X_\tau.$$

U_s is still a supermartingale. In the case where Q is a *tree*, Haggstrom [14], following Snell [22], obtained sufficient conditions for the existence of an optimal tactic in terms of a comparison of X_t and U_t. (Haggstrom does not have the notion of a tactic, but it is easy to see that in the case of a tree, his "control variable" defines a tactic.) In a rather weak sense, Haggstrom's theorem extends to a more general order, because for the purpose of this theorem it is possible to map a general Q on a tree. Such a map, however, is not one-to-one, and it destroys the properties of σ-algebras considered in the remainder of this article, in particular the "conditional qualitative independence."

2. Conditions for Representing Stopping Rules as Tactics

Q is again a countable locally finite set. We study stopping rules and tactics relative to a fixed stochastic basis $(\mathfrak{F}_t, t \in Q)$.

Two σ-algebras $\mathcal{A}, \mathcal{B}$ are said to be *conditionally qualitatively independent* with respect to a σ-algebra $\mathcal{C}$, in symbols $\mathcal{A} \perp \mathcal{B} \mid \mathcal{C}$, if for any sets $A \in \mathcal{A}$, $B \in \mathcal{B}$, $\{P^{\mathcal{C}}A > 0\} \cap \{P^{\mathcal{C}}B > 0\} \subset \{P^{\mathcal{C}}(A \cap B) > 0\}$.

Given $s \in Q$, let the set of direct successors of s be $D_s = \{s^1, s^2, ...\}$.

Let $M(s) = \{t \in Q: t > s\}$. A family of disjoint subsets of Q, $\mathscr{E}(s) = \mathscr{E} = \{E_1, E_2, ...\}$, is called *complete* (for s) if

(i) $\bigcup E_i \subset M(s)$,

(ii) $s^j \in E_j$ for every j,

and

(iii) for every j, $M(s) \backslash \bigcup_{i \neq j} E_i \subset L(s^j)$.

The family $\mathscr{E}$ is said to satisfy the condition CQI (conditional qualitative independence) if for any i, j, $i \neq j$, any $t \in E_i$, $u \in E_j$, one has that $\mathfrak{F}_t \perp \mathfrak{F}_u \mid \mathfrak{F}_s$. *The stochastic basis* $(\mathfrak{F}_t, t \in Q)$ is said to *satisfy the condition* CQI, if for each s there is a complete family $\mathscr{E}(s)$ satisfying CQI.

In order to explain the meaning of the condition CQI, we describe important particular cases in which CQI may hold.

(1) Q is a *plane rectangle*. By this we mean that each element s of Q is a pair of positive integers, $s = (s_1, s_2)$, with the partial order defined by $s \leqslant t \Leftrightarrow (s_1 \leqslant t_1$ and $s_2 \leqslant t_2)$. The smallest element is $p = (1, 1)$. For some $z = (z_1, z_2)$ with $z_1 \leqslant \infty$, $z_2 \leqslant \infty$, $Q = \{t = (t_1, t_2): t_1 < z_1, t_2 < z_2\}$.

One may choose $E_1 = \{t: t_1 > s_1, t_2 = s_2\}$, $E_2 = \{t : t_1 = s_1, t_2 > s_2\}$. The condition CQI will obviously hold if the σ-algebras $\mathfrak{F}_s$ are generated by

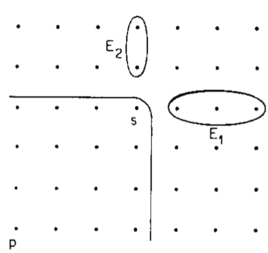

FIGURE 1

independent random variables Y_r, $r \leqslant s$. Hence it will be in particular possible to consider random variables X_s which are sums or averages of independent random variables Y_r, $r \leqslant s$.

(2) Q is a *tree*. This means that if $D_s = \{s^1, s^2, ...\}$, and $j \neq k$, then no element in $L(s^j)$ is comparable with any element in $L(s^k)$. Set $E_j = L(s^j)$. Again CQI holds if each $\mathfrak{F}_s$ is generated by independent Y_r, $r \leqslant s$.

THEOREM 2.1. *Suppose that the stochastic basis* $(\mathfrak{F}_t, t \in Q)$ *satisfies the conditional qualitative independence condition* CQI. *Then every stopping time is given by a tactic*; *more precisely, for each* $\tau \in \bar{\Sigma}$ *there exists a tactic* $\mathscr{H}$ *such that* $\tau = \tau_{\mathscr{H}}$.

Proof. Let $\tau \in \bar{\Sigma}$. Fix s; let $D_s = \{s^1, s^2, ...\}$, and let $\mathscr{E} = \mathscr{E}(s) = \{E_1, E_2, ...\}$ be a complete family with $s^i \in E_i$. Let t and u be two elements in the range of τ with $t \in E_i$, $u \in E_j$, $i \neq j$. The sets $A = \{\tau = t\}$ and $B = \{\tau = u\}$ are disjoint; hence $P^{\mathfrak{F}_s}(A \cap B) = 0$. By CQI, for each ω either $P^{\mathfrak{F}_s}A(\omega) = 0$, or $P^{\mathfrak{F}_s}B(\omega) = 0$. Let $F_i = \{\omega : P^{\mathfrak{F}_s}(\tau \in E_i) > 0\}$; it follows that the sets F_i are disjoint. Define $H_{ss} = \{\tau = s\}$, $H_{ss^i} = F_i$ for $i \geqslant 2$, and

$$H_{ss^1} = F_1 \cup \left(H_{ss} \cup \bigcup_{i \geqslant 2} F_i\right)^c.$$

It will be easy to see that $\mathscr{H} = \{H_{st}\}$ is a tactic. Now suppose $P(\tau = t) > 0$, let $\omega \in \{\tau = t\}$, $s < t$. Then $\omega \notin H_{ss}$. Suppose that at ω, we "go to s^j"; i.e., $\omega \in H_{ss^j}$. If $t = s^j$, $\omega \in H_{s^j s^j}$, and $\tau_{\mathscr{H}}(\omega) = s^j = t$. If $t \neq s^j$, by the construction of sets H_{ss^i}, we know that $t \notin \bigcup_{i \neq j} E_i$. Hence by the assumption (iii) of the condition CQI, $t \in L(s^j)$. The argument is now resumed with s^j replacing s. After finitely many steps, one obtains $s^j = t$. It follows that $\{\tau = t\} = \{\tau_{\mathscr{H}} = t\}$. Similarly, $\{\tau = \infty\} = \{\tau_{\mathscr{H}} = \infty\}$. ∎

Remark. There is some arbitrariness in the procedure used in the proof, because if the values of $\tau > s$ do not lie in any of the E_i's , we choose the tactic that "goes from s to s^1," while insted of s^1 any other direct successor of s could be chosen. Thus in general many tactics correspond to the same stopping time. To avoid this, we could assume that one always chooses s^1, the elements of D_s being enumerated according to an auxiliary total order $\prec$ defined on Q.

If $s \in Q$, the set where a tactic $\mathscr{H}$ *passes through* s is by definition

$$P_{\mathscr{H}}(s) = P(s) = \bigcup \bigcap_{i=0}^{n} H_{t(i)t(i+1)},$$

where the union is over all finite sequences $p = t(0) < t(1) < \cdots < t(n) \leqslant t(n+1)$, with $t(i+1) \in D_{t(i)}$, $i = 0, 1, ..., n-1$; $t(n) = s$, $t(n+1) \in \{s\} \cup D_s$.

The set $S_{\mathscr{H}}(s) = S(s)$, where the tactic $\mathscr{H}$ *stops at s* is $P(s) \cap H_{ss}$.

Theorem 2.1 admits the following partial converse:

THEOREM 2.2. *Let Q be a plane rectangle. Let $E_1 = \{t\colon t_1 > s_1,\ t_2 = s_2\}$, $E_2 = \{t\colon t_1 = s_1,\ t_2 > s_2\}$. If $\mathscr{E} = (E_1, E_2)$ does not satisfy the condition CQI, then there is a stopping rule $\sigma \in \Sigma$ such that there exists no tactic $\mathscr{H}$ with $\sigma = \tau_{\mathscr{H}}$.*

Proof. There are elements $t = (t_1, t_2) \in E_1$, $u = (u_1, u_2) \in E_2$, and sets $A \in \mathfrak{F}_t$, $B \in \mathfrak{F}_u$ such that on a set $D \in \mathfrak{F}_s$ one has $P^{\mathfrak{F}_s}A > 0$, $P^{\mathfrak{F}_s}B > 0$, but on a non-null $\mathfrak{F}_s$-measurable subset E of D, $P^{\mathfrak{F}_s}(A \cap B) = 0$.

Then $P(A \cap E) = \int_E P^{\mathfrak{F}_s}A > 0$, and also $P(B \cap E) > 0$, but the sets $A \cap E$ and $B \cap E$ are disjoint. Define $\sigma \in \Sigma$ by $\sigma = t$ on $A \cap E$, $\sigma = u$ on $B \cap E$, $\sigma = v = (t_1, u_2)$ elsewhere. Assume $\sigma = \tau_{\mathscr{H}}$. Since $P^{\mathfrak{F}_s}(\sigma = t) > 0$ on E, $P(s) \cap E \subset H_{ss^1}$. Similarly $P(s) \cap E \subset H_{ss^2}$. If $P[P(s)] = 1$, this leads to a contradiction, because the sets H_{ss^1} and H_{ss^2} are disjoint, and $P[P(s) \cap E] = P(E) > 0$.

It remains to show that the assumption $P[P(s)] = 1$ is not a loss of generality. This from the following lemma:

LEMMA 2.3. *Let $\mathscr{H}'$ be a tactic on a plane rectangle with $\tau_{\mathscr{H}'} = \sigma$, where σ is defined as above. Then there exists another tactic $\mathscr{H}$ with $\tau_{\mathscr{H}} = \sigma$ and such that $P_{\mathscr{H}}(s) = \Omega$.*

Proof. $P_{\mathscr{H}}(s)$ is set where the tactic $\mathscr{H}$ passes through s. Let the first $s_1 + s_2 - 1$ elements of the sequence $t(i)$ in the definition of $\tau_{\mathscr{H}}$ be: (1, 1), (2, 1),..., $(s_1, 1)$, $(s_1, 2)$,..., (s_1, s_2). Recall that s^1 and s^2 are the first elements of E_1 and Z_2, respectively. Let

$$H_{ss^2} = \bigcup_{x \leqslant s_1} P_{\mathscr{H}'}(x, s_2) \cap H'_{(x, s_2)(x, s_2+1)},$$

$$H_{ss^1} = \bigcup_{y \leqslant s_2} P_{\mathscr{H}'}(s_1, y) \cap H'_{(s_1, y)(s_1+1, y)}.$$

Then the definition of $\mathscr{H}$ can be continued so that $\tau_{\mathscr{H}} = \sigma$. Informally, the continuation can be described by saying that on $H_{ss^1} \cap S_{\mathscr{H}'}(t)$ the tactic $\mathscr{H}$ proceeds to t and stops there; on $H_{ss^1} \cap S_{\mathscr{H}'}(v)$ $\mathscr{H}$ proceeds to t, then continues on to v and stops on v; on $H_{ss^2} \cap S_{\mathscr{H}'}(u)$ the tactic $\mathscr{H}$ proceeds to u and stops there; on $H_{ss^2} \cap S_{\mathscr{H}'}(v)$ the tactic $\mathscr{H}$ proceeds to u, then continues on to v and stops at v. Clearly, $P_{\mathscr{H}}(s) = \Omega$ and $\tau_{\mathscr{H}} = \sigma$. ∎

In what follows, $Q = \mathbb{N}^d$ denotes the d-dimensional discrete space with the coordinatewise (partial) order: $s = (s_1, ..., s_d) \leqslant t = (t_1, ..., t_d)$ if and only if $s_1 \leqslant t_1, ..., s_d \leqslant t_d$. The following example shows that if $d = 3$, then there are stopping rules not given by tactics, even though the σ-algebras $\mathfrak{F}_t$ are each generated by independent random variables Z_s, $s \leqslant t$.

EXAMPLE OF A STOPPING RULE IN $\mathbb{N}^3$, WHICH CANNOT BE REPLACED BY A TACTIC. $\Omega = [0, 1]^3$ with P being the three-dimensional Lebesgue measure. Z_1, Z_2, Z_3 are coordinate variables. Put

$$C_1 = \{z \in \Omega: z_2 \leqslant \tfrac{1}{2}, z_3 \leqslant \tfrac{1}{2}\},$$
$$C_2 = \{z \in \Omega: z_1 > \tfrac{1}{2}, z_3 > \tfrac{1}{2}\},$$
$$C_3 = \{z \in \Omega: z_1 \leqslant \tfrac{1}{2}, z_2 > \tfrac{1}{2}\}.$$

These sets are disjoint, each of measure $\frac{1}{4}$. Now consider the set $Q = \{1, 2\}^3$ with coordinatewise order and set

$$\mathfrak{F}_{(1,1,1)} = \{\varnothing, \Omega\}, \mathfrak{F}_{(1,1,2)} = \sigma(Z_3) = \text{the } \sigma\text{-algebra generated by } Z_3,$$
$$\mathfrak{F}_{(1,2,1)} = \sigma(Z_2), \mathfrak{F}_{(2,1,1)} = \sigma(Z_1),$$
$$\mathfrak{F}_{(1,2,2)} = \sigma(Z_2, Z_3), \mathfrak{F}_{(2,1,2)} = \sigma(Z_1, Z_3), \mathfrak{F}_{(2,2,1)} = \sigma(Z_1, Z_2),$$
$$\mathfrak{F}_{(2,2,2)} = \sigma(Z_1, Z_2, Z_3).$$

Put

$$\begin{aligned} \tau(z) &= (1, 2, 2) && \text{for} \quad z \in C_1 \\ &= (2, 1, 2) && \text{for} \quad z \in C_2 \\ &= (2, 2, 1) && \text{for} \quad z \in C_3 \\ &= (2, 2, 2) && \text{elsewhere.} \end{aligned}$$

Put

$$X_{(1,1,1)} = X_{(1,1,2)} = X_{(1,2,1)} = X_{(2,1,1)} = X_{(2,2,2)} \equiv 0,$$
$$X_{(1,2,2)} = 1_{C_1}, X_{(2,1,2)} = 1_{C_2}, X_{(2,2,1)} = 1_{C_3}.$$

Then $EX_\tau = \frac{3}{4}$.

Claim. There is no $\tau_{\mathscr{T}}$ generated by a tactic which gives $EX_{\tau_{\mathscr{T}}} \geqslant \frac{3}{4}$.

Proof. $\mathfrak{F}_{(1,1,1)}$ is trivial. Therefore we would have to stop in (1, 1, 1) on all of Ω (which is bad), or decide for *one* of the points (1, 1, 2), (1, 2, 1), (2, 1, 1) to be our $t(1)$ on *all* of Ω. In each case passing through one of the points (1, 2, 2), (2, 2, 1), (2, 1, 2) is ruled out and we cannot stop in that point any more. The integral of $\sup X_t$ over the remaining points is $\frac{1}{2} < \frac{3}{4}$. ∎

3. Linear Embedding of Tactics

We shall now define a mapping which provides a general method for extending results on stopping rules for independent identically distributed processes $Y_1, Y_2,\dots$ to independent identically distributed processes indexed by a finite or countable locally finite index set Q and the stopping times obtained from tactics. For example the theorem of Dvoretzky and Davis on the existence of optimal stopping times for $X_n = n^{-1} \sum_{i=1}^n Y_i$ will immediately extend to tactics for processes $\{Y_t, t \in Q\}$ and in particular to all stopping rules in the case $Q = \mathbb{N}^2$. Other applications will be an extension of Wald's [23] identity in the form given to it by Robbins and Samuel [19], and of the corresponding result of Blackwell and Girshick [1] on second moments. The method also provides an answer to a problem of Cairoli and Gabriel [5] on stopping rules and the class $L \log L$ for dimension 2. These applications will be treated in the subsequent section.

Let $(Y_t, t \in Q)$ be random variables taking values in a measurable space $(E, \mathscr{E})$, and defined on $(\Omega, \mathfrak{F}, P)$. Let $\mathfrak{F}_t$ be the σ-algebra $\sigma(Y_s, s \leqslant t)$ generated by $\{Y_s, s \leqslant t\}$. As always when we deal with tactics we assume the existence of a $p \in Q$ with $p \leqslant q$ for all $q \in Q$.

To avoid technicalities at the beginning, we assume at first that Ω is the space E^Q of mappings $\omega: Q \to E$, and that Y_t is the tth coordinate variable $Y_t(\omega) = \omega(t) = \omega_t$.

Let us first look at the case of a finite index set Q. $|Q|$ is the cardinality of Q. Let $\mathscr{H} = \{H_{s,t}, s \in Q, t \in \{s\} \cup D_s\}$ be a tactic. We associate with $\mathscr{H}$ a random map $\varphi = \{\varphi_\omega, \omega \in \Omega\}$, i.e., a family of maps $\varphi_\omega: Q \to \{1, 2,\dots, |Q|\}$ as follows: We know that there exists a finite sequence

$$p = t(0) < t(1) < t(2) < \cdots < t(n) = t(n+1),$$

with $n \geqslant 0$ and the $t(k)$'s depending on ω, such that $t(i) \in D_{t(i-1)}$ $(i = 1,\dots, n)$, $\omega \in H_{t(i-1)t(i)}$ $(i = 1,\dots, n+1)$, and $\tau_{\mathscr{H}}(\omega) = t(n)$. (As Q is finite, $\tau_{\mathscr{H}}(\omega) = \infty$ cannot occur.) We may express the dependence on ω by writing $n = n_\omega$ and $t(k) = t_\omega(k)$ when necessary. Put $\varphi_\omega(p) = 1$ for all $\omega \in \Omega$. Write $K(t) = \{s \in Q: s \leqslant t\}$. $|t| = |K(t)| = \operatorname{card} K(t)$. For any k with $0 \leqslant k \leqslant n = n(\omega)$, φ_ω will be an invertible map of the set

$$V(k) = V_\omega(k) = K(t(k+1)) \backslash K(t(k)$$

onto the interval $\{i \in \mathbb{N}: |t(k)| < i \leqslant |t(k+1)|\}$. We require that $\varphi_\omega(t(k+1)) = |t(k+1)|$. This does not yet determine φ_ω uniquely. To eliminate this ambiguity we fix an arbitrary enumeration $\{q_1, q_2,\dots\}$ of Q. This determines a total order $\prec$ by $q_1 \prec q_2 \prec \cdots$. Now φ_ω is uniquely determined on $V_\omega(k)$ if we ask that φ_ω be $\prec$-isotonic on $V_\omega(k) \backslash \{t(k+1)\}$, i.e.,

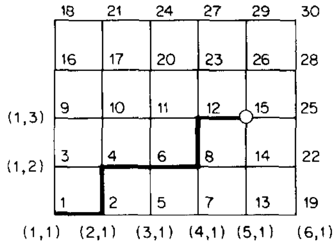

FIGURE 2

$u, v \in V_\omega(k)\backslash\{t(k+1)\}$, $u \prec v$ shall imply $\varphi_\omega(u) < \varphi_\omega(v)$. Finally φ_ω maps $K(t(n))^c$ in an $\prec$-isotonic way onto $\{|t(n)|+1,..., |Q|\}$.

In Fig. 2 we try to illustrate the map φ_ω for $Q = \{1, 2,..., 6\} \times \{1, 2,..., 5\}$, assuming $(1,1) \prec (2,1) \prec (1,2) \prec (3,1) \prec (2,2) \prec (1,3) \prec (4,1) \prec \cdots$ and $(i_1, i_2) \leqslant (j_1, j_2)$ iff $i_1 \leqslant j_1$ and $i_2 \leqslant j_2$. For some ω we may have $(1,1) = p = t(0) < t(1) = (2,1) < t(2) = (2,2) < t(3) = (3,2) < t(4) = (4,2) < t(5) = (4,3) < t(6) = (5,3) = t(7)$, because ω belongs to the sets $H_{t(i-1)t(i)}$. The "path" $t(0) < \cdots < t(n) \leqslant t(n+1)$ is rendered in heavy print. The obvious indices $(i, j) \in Q$ are given only for a few elements of Q. The integers next to the elements (i, j) are $\varphi_\omega(i, j)$. Stopping occurs at $(5, 3)$.

The random map $\varphi = \{\varphi_\omega\}$ induces a map $\Psi: E^Q \to E^{|Q|}$ as follows: If ρ_ω is the inverse of φ_ω put for $\omega = (\omega_t, t \in Q) \in E^Q$

$$\Psi(\omega) = (\eta_1, \eta_2,..., \eta_{|Q|})$$

with $\eta_k = \omega_{\rho_\omega(k)}$. What we have done is the following: Depending on ω the indices $t \in Q$ have received new names $\varphi_\omega(t)$. So $\Psi(\omega)$ is just a way of writing down the ω_t in the order of the new names of the indices of ω.

Let us check that Ψ is a bijective map of E^Q onto $E^{|Q|}$: It is clear that $\eta_1 = \omega_p$. Because of $\mathfrak{F}_p = \sigma(Y_p)$, the coordinate ω_p determines $t_\omega(1)$. If $t_\omega(1) = p$, i.e., the tactic stops at p, $\eta_2, \eta_3,...$ are just the other coordinates of ω in the order of $\prec$, and we know ω. If $t_\omega(1) = s \in D_p$, this determines the map φ_ω on $K(s)$ and therefore the first $|t_\omega(1)|$ coorinates of η now determine ω_t for $t \leqslant t_\omega(1)$, and therefore $t_\omega(2)$. This argument can be continued.

If $|Q| = \infty$, we have to distinguish two cases. If $\tau_{\mathscr{T}}(\omega)$ is finite, we get $\Psi(\omega)$ as above, and again $\Psi(\omega)$ determines ω. If $\tau_{\mathscr{T}}(\omega) = \infty$, we have an infinite sequence $p = t(0) < t(1) < t(2) < \cdots$. We must consider the case when the union W_ω of the sets $K(t_\omega(k))$ is not all of Q. φ_ω is then defined only on W_ω. Nevertheless, for any k, $\rho_\omega(k) \in W_\omega$ is well defined. If $Y_t(\omega)$ and $Y_t(\omega')$ agree for all t in W_ω, but differ for some other coordinates, we still get $\Psi(\omega) = \Psi(\omega')$. Ψ is still a mapping onto $E^{|Q|}$, but need not be invertible. Yet, we can define a random variable $\sigma_{\mathscr{H}}$ on $E^{|Q|}$ by

$$\sigma_{\mathscr{H}} = |\tau_{\mathscr{H}} \circ \Psi^{-1}|$$

in a unique way: If $\omega, \omega' \in E^Q$ have the property that $\Psi(\omega) = \Psi(\omega') = \eta = (\eta_1, \eta_2, \ldots) \in E^{|Q|}$, then $\tau_{\mathscr{H}}(\omega') = \tau_{\mathscr{H}}(\omega)$ whether the value is finite or not. $\sigma_{\mathscr{H}}$ will be called the *embedding of* $\tau_{\mathscr{H}}$ *in the line.*

Let $\Omega' = E^{|Q|}$ and let $\mathfrak{F}'$ be the σ-algebra generated by the coordinate variables Y'_k $(k \geqslant 1)$ in Ω', and $\mathfrak{F}'_k = \sigma(Y'_i, i \leqslant k)$.

THEOREM 3.1. *Let $Q, \Omega = E^Q$, etc., be as above and let $\mathscr{H}$ be a tactic for the stochastic basis $(\mathfrak{F}_t,\ t \in Q)$ generated by the coordinate variables Y_t. Then $\sigma_{\mathscr{H}}$ is a stopping time for $\mathfrak{F}'_k$ $(k \geqslant 1)$. If the Y_t are independent and identically distributed under P in Ω and $P' = P \circ \Psi^{-1}$, then the Y'_k are independent and identically distributed with the same distribution as we Y_t.*

Proof. We have to show that $\{\eta \in E^{|Q|}: \sigma_{\mathscr{H}}(\eta) = k\}$ depends only on the coordinate variables $Y'_1, \ldots, Y'_k$. If $\sigma_{\mathscr{H}}(\eta) = 1$, then $\tau_{\mathscr{H}}(\Psi^{-1}(\eta)) = p$. This means that $\Psi^{-1}(\eta) \in H_{pp}$. As $H_{pp} \in \sigma(Y_p)$, there exists an $H^*_{pp} = Y_p^{-1}(H^*_{pp})$. Thus $\sigma_{\mathscr{H}}(\eta) = 1$ is equivalent with $Y_p(\Psi^{-1}(\eta)) \in H^*_{pp}$. Now $Y'_1(\eta) = Y_p(\Psi^{-1}(\eta))$ implies $\{\eta: \sigma_{\mathscr{H}}(\eta) = 1\} \in \mathfrak{F}'_1$.

Now take any $\eta = (\eta_1, \eta_2, \ldots)$ with $\sigma_{\mathscr{H}}(\eta) = k$, and look at some $\tilde{\eta} = (\tilde{\eta}_1, \tilde{\eta}_2, \ldots)$ such that $\eta_1 = \tilde{\eta}_1, \ldots, \eta_k = \tilde{\eta}_k$. We want to show $\sigma_{\mathscr{H}}(\tilde{\eta}) = k$. We may assume $k > 1$.

Let $\omega, \tilde{\omega}$ be elements of Ω with $\Psi(\omega) = \eta$ and $\Psi(\tilde{\omega}) = \tilde{\eta}$. $\omega_p = \eta_1 = \tilde{\eta}_1 = \tilde{\omega}_p$ implies that ω and $\tilde{\omega}$ lie in the same H_{ps}, say, in $H_{pt(1)}$. But this $t(1)$ must then be the $t_\omega(1)$ and also the $t_{\tilde{\omega}}(1)$, so that $\rho_\omega(l) = \rho_{\tilde{\omega}}(l)$ for $l = 1, \ldots, |t(1)|$. It is impossible that $k < |t(1)|$ because $\tau_{\mathscr{H}}(\omega)$ must be an element $t_\omega(n)$. I $k = |t(1)|$ then $\tau_{\mathscr{H}}(\omega) = t(1)$. In this case, because of $\rho_\omega(l) = \rho_{\tilde{\omega}}(l)$ and $\eta_l = \tilde{\eta}_l$, we get $\omega_s = \tilde{\omega}_s$ for $s \leqslant t(1)$. Then also $\tau_{\mathscr{H}}(\tilde{\omega}) = t(1)$ and $\sigma_{\mathscr{H}}(\tilde{\eta}) = k$.

If $k > |t(1)|$, $\omega_s = \tilde{\omega}_s$ $(s \leqslant t(1))$ implies that also $\sigma_{\mathscr{H}}(\tilde{\omega}) > |t(1)|$ and $t_\omega(2) = t_{\tilde{\omega}}(2)$. We can now argue as before. After at most k steps we must have $|t_\omega(n)| = k$ and $|t_\omega(n)| = |t_{\tilde{\omega}}(n)|$ and then $\sigma_{\mathscr{H}}(\tilde{\omega}) = k$.

This implies that $\{\eta: \sigma_{\mathscr{H}}(\eta) = k\}$ depends only on $Y'_1, \ldots, Y'_k$. The fact that the set is measurable in the σ-algebra generated by $Y'_1, \ldots, Y'_k$ follows from the fact that all the constructions of φ and Ψ were carried out in a measurable way.

We now assume that the Y_t $(t \in Q)$ are independent and identically distributed. Y'_1 has the same distribution as Y_p beecause for any $A_1 \in \mathscr{E}$, $\Psi^{-1}(\{Y'_1 \in A_1\}) = \{Y_p \in A_1\}$.

Now consider $A_1, A_2 \in \mathscr{E}$. For $s \in D_p$ let $t(s, 1)$ be the first element of

$K(s)\backslash\{p\}$ and $u(p, 1)$ the first element of $Q\backslash\{p\}$ in the $\prec$ order. Then for $\omega \in H_{ps}$ we have $\varphi_\omega(t(s, 1)) = 2$ and for $\omega \in H_{pp}$ we have $\varphi_\omega(u(p, 1)) = 2$. Hence

$$\begin{aligned}
&\Psi^{-1}(\{Y_1' \in A_1, Y_2' \in A_2\})\\
&= \left(\bigcup_{s \in D_p} H_{ps} \cap \Psi^{-1}(\{Y_1' \in A_1, Y_2' \in A_2\})\right)\\
&\quad \cup (H_{pp} \cap \Psi^{-1}\{Y_1' \in A_1, Y_2' \in A_2\})\\
&= \left(\bigcup_{s \in D_p} H_{ps} \cap \{Y_p \in A_1, Y_{t(s,1)} \in A_2\}\right)\\
&\quad \cup \left(H_{pp} \cap \{Y_p \in A_1, Y_{u(p,1)} \in A_2\}\right).
\end{aligned}$$

It follows that

$$\begin{aligned}
&P'(Y_1' \in A_1, Y_2' \in A_2)\\
&= \sum_{s \in D_p} P(H_{ps} \cap \{Y_p \in A_1, Y_{t(s,1)} \in A_2\})\\
&\quad + P(H_{pp} \cap \{Y_p \in A_1, Y_{u(p,1)} \in A_2\}).
\end{aligned}$$

Now H_{ps}, H_{pp}, and Y_p are independent of $Y_{t(s,1)}$ and $Y_{u(p,1)}$. We can therefore pull out the factors $P(Y_{t(s,1)} \in A_2)$ and $P(Y_{u(p,1)} \in A_2)$. As they both agree with $P(Y_p \in A_2)$, we find that

$$P'(Y_1' \in A_1, Y_2' \in A_2) = P(Y_p \in A_1)\, P(Y_p \in A_2).$$

The argument can be continued: For the computation of $P'(Y_1' \in A_1, ..., Y_m' \in A_m)$ the space is split into subsets such that on each subset $Y_m' \circ \Psi$ is the same Y_t. This can be done in such a way that the subset W_{mt} on which $Y_m' \circ \Psi$ is given by Y_t is independent of Y_t and also $W_{mt} \cap \Psi^{-1}(\{Y_1' \in A_1, ..., Y_{m-1}' \in A_{m-1}\})$ is independent of Y_t. Hence

$$\begin{aligned}
&P'(Y_1' \in A_1, ..., Y_m' \in A_m)\\
&\quad = P\left(\Psi^{-1}(\{Y_1' \in A_1, ..., Y_{m-1}' \in A_{m-1}\}) \cap \left(\bigcup_t (W_{mt} \cap \Psi^{-1}(Y_m' \in A_m))\right)\right)\\
&\quad = \sum_{t \in Q} P(\Psi^{-1}(\{Y_1' \in A_1, ..., Y_{m-1}' \in A_{m-1}\}) \cap W_{mt} \cap \{Y_t \in A_m\})\\
&\quad = \sum_{t \in Q} P(\Psi^{-1}(\{Y_1' \in A_1, ..., Y_{m-1}' \in A_{m-1}\}) \cap W_{mt})\, P(Y_p \in A_m)\\
&\quad = P'(Y_1' \in A_1, ..., Y_{m-1}' \in A_{m-1})\, P(Y_p \in A_m).
\end{aligned}$$

This proves the theorem. The explicit construction of the set W_{mt} in terms of the possible sequences $p = t(0) < t(1) < \cdots$ would be clumsy, but it is clear from the construction of Ψ that it can be done. ∎

In applications of the above theorem to the theory of optimal stopping one must also specify a reward X_t for stopping at t ($t \in Q$).

Let $f = \{f_n, n \in \mathbb{N}\}$ be a sequence of measurable mappings $f_n: E_n \to \mathbb{R}^1$, symmetric in all variables other than the first one; i.e., for all $n \in \mathbb{N}$, all permutations $(\pi_2, \pi_3, \ldots, \pi_n)$ of $(2, 3, \ldots, n)$ and all $(e_1, \ldots, e_n) \in E^n$, we assume

$$f_n(e_1, e_2, \ldots, e_n) = f_n(e_1, e_{\pi_2}, \ldots, e_{\pi_n}).$$

The range of f_n may also be a Banach space. For example, if $E = \mathbb{R}^1$ we can take $f_n(e_1, \ldots, e_n) = e_1$, or $f_n(e_1, \ldots, e_n) = n^{-1}(e_1 + \cdots + e_n)$, or $f_n = \prod_{i=1}^n e_i$. Using the family $(Y_t, t \in Q)$ and $f = \{f_n\}$ we can define a reward process X_t by

$$X_t = f_{|t|}(Y_t, Y_s: s < t).$$

We do not have to give an order in which the Y_s with $s < t$ appear, because of the symmetry with respect to the variables $e_2, \ldots, e_n$. So with the first choice of f above we would get $X_t = Y_t$ and with the second choice $X_t = |t|^{-1} \sum_{s \leqslant t} Y_s$. (We hope it does not disturb the reader that the "last" random variable Y_t appears at the location of the first variable e_1 of $f_{|t|}(e_1, \ldots)$. It is sometimes more important than the other Y_s $(s < t)$.)

Put $X'_k = f_k(Y'_k, Y'_i: 1 \leqslant i < k)$. We are interested in $X_{\tau_{\mathscr{Y}}}$ and $X'_{\sigma_{\mathscr{Y}}}$. As $\tau_{\mathscr{Y}}$ and $\sigma_{\mathscr{Y}}$ may be infinite we just define $X_\infty = X'_\infty = c_\infty$, where c_∞ is an arbitrary constant, e.g., $c_\infty = 0$.

THEOREM 3.2. *If the reward functions X_t and X'_k are given by an $f = \{f_n\}$ as above, we have $X_{\tau_{\mathscr{Y}}} = X'_{\sigma_{\mathscr{Y}}} \circ \Psi$, so that the distribution of $X_{\tau_{\mathscr{Y}}}$ under P coincides with the distribution of $X'_{\sigma_{\mathscr{Y}}}$ under $P' = P \circ \Psi^{-1}$. In particular $E_P X_{\tau_{\mathscr{Y}}} = E_{P'} X'_{\sigma_{\mathscr{Y}}}$ when one and hence both expectations exist.*

Remark. The interest of this theorem lies in the case where the Y_t are independent identically distributed, because then the embedding theorem tells us that the Y'_k are also independent identically distributed, with the same distribution.

Proof of Theorem. First consider ω with $\tau_{\mathscr{Y}}(\omega) < \infty$. $X_{\tau_{\mathscr{Y}}}(\omega) = f_{|\tau_{\mathscr{Y}}(\omega)|}(Y_{\tau_{\mathscr{Y}}}(\omega); Y_s: s < \tau_{\mathscr{Y}}(\omega))$. The coordinates $Y_s(\omega)$ with $s < \tau_{\mathscr{Y}}(\omega)$ all appear (by the construction of Ψ) with the correct frequencies among the coordinates $Y'_k(\Psi(\omega))$ $(k < |\tau_{\mathscr{Y}}(\omega)|)$. From $|\tau_{\mathscr{Y}}(\omega)| = \sigma_{\mathscr{Y}}(\Psi(\omega))$ and $Y_{\tau_{\mathscr{Y}}}(\omega) = Y'_{\sigma_{\mathscr{Y}}}(\Psi(\omega))$ we therefore get

$$X_{\tau_{\mathscr{Y}}}(\omega) = f_{\sigma_{\mathscr{Y}}(\Psi(\omega))}(Y'_{\sigma_{\mathscr{Y}}}(\Psi(\omega)), Y'_k(\Psi(\omega)): k < \sigma_{\mathscr{Y}}(\Psi(\omega))) = X'_{\sigma_{\mathscr{Y}}}(\Psi(\omega)).$$

If $\tau_{\mathscr{H}}(\omega)=\infty$ then also $\sigma_{\mathscr{H}}(\Psi(\omega))=\infty$, and the identity follows from $X_\infty \equiv X'_\infty \equiv c_\infty$. ∎

The Passage to General Probability Spaces. Above we have made the assumption $\Omega = E^Q$ only for convenience's sake. If $(\Omega^*, \mathfrak{F}^*, P^*)$ is an abstract probability space, $Y^* = (Y_t^*, t \in Q)$ a family of E-valued random variables, and $\mathscr{H}^* = \{H_{st}^*;\ s \in Q,\ t \in \{s\} \cup D_s\}$ a tactic for the stochastic basis $\mathfrak{F}_t^* = \sigma(Y_s^*, s \leqslant t)$ we can obtain an embedding by a reduction to the case $\Omega = E^Q$. The sets H_{st}^* are of the form $Y^{*-1}H_{st}$ for a tactic $\mathscr{H}$ in E^Q. Thus we can apply the above result with $P = P^* Y^{*-1}$. With $X_t^* = f_{|t|}(Y_t^*, Y_s^*, s < t)$ one gets $X_{\tau_{\mathscr{H}^*}}^* = X'_{\sigma_{\mathscr{H}}} \circ \Psi \circ Y^*$ and $E_{P^*} X_{\tau_{\mathscr{H}^*}}^* = E_{P'} X'_{\sigma_{\mathscr{H}}}$. In particular, when the Y_t^* are independent identically distributed, P' is known, and we again have a reduction of stopping problems with index set Q to problems with index set $\mathbb{N}$. In Section 5 on ordered stopping times, we shall see that this reduction is more difficult for stopping times other than tactics.

4. Applications of the Linear Embedding Theorem

Stopping rules and tactics in this section will always be with respect to $\mathfrak{F}_t = \sigma(Y_s, s \leqslant t)$.

(a) Existence of optimal stopping rules and tactics: Let Y_n $(n \in \mathbb{N})$ be independent identically distributed random variables with $EY_n = 0$ and $\operatorname{Var} Y_n < \infty$. Extending a result of Chow and Robbins [7] on the coin-tossing case, Dvoretzky [12] has proved the existence of a finite optimal stopping rule σ for $X_n = n^{-1} \sum_{i=1}^n Y_i$. Davis [11] weakened the assumption $\operatorname{Var} Y_n < \infty$ to $E|Y_n|^p < \infty$ for some $p > 1$.

Now consider $Q = \mathbb{N}^d$ and a family $(Y_t, t \in Q)$ with $EY_t = 0$, $\operatorname{Var} Y_t = 1$. We may view $\mathbb{N}$ as the subset $\{t = (t_1, ..., t_d) \in \mathbb{N}^d\colon t_2 = t_3 = \cdots = t_d = 1\}$. Thus σ can be viewed as a stopping rule (given by a tactic) taking values in Q. The embedding theorem tells us that σ is optimal for the reward process $X_t = |t|^{-1} \sum_{s \leqslant t} Y_s$ among all tactics stopping in Q, and not only in $\mathbb{N}$. In the case $d = 2$, Theorem 2.1 tells us that every stopping rule is given by a tactic, so that σ is then optimal among all stopping rules. Thus the several parameter case can be reduced to the 1-parameter case treated by Dvoretzky and Davis.

This is in fact a general device. Whenever there is an optimal σ for $X_n = f_n(Y_n, Y_1, ..., Y_{n-1})$ with independent identically distributed random variables other than the first, then the same stopping rule is also optimal among the tactics for the corresponding process $(X_t, t \in Q = \mathbb{N}^d)$. For example, with $f_n(e_1, ..., e_n) = n^{-1} e_1$ we can extend results of Chow and Robbins [7] and of Chow and Dvoretzky [6] on stopping $X_n = n^{-1} Y_n$. (See Chap. V, Sections 6–8 of [8].)

The argument also can be used for trees Q if Q contains an infinite sequence $q^1 < q^2 < \cdots$, or if the length of all finite sequences $q^1 < q^2 < \cdots < q^n$ in Q is bounded.

(b) A problem of Cairoli and Gabriel: The embedding can also be useful when the existence of an optimal stopping rule is not known in the 1-parameter case. We shall now apply it to give a positive answer to a problem of Cairoli and Gabriel [5] in the case $d = 2$.

Let $(Y_t,\ t \in Q = \mathbb{N}^d)$ again be independent identically distributed random variables, and $X_t = |t|^{-1} \sum_{s \leqslant t} Y_s$. Burkholder [4] has shown for $d = 1$ and Gabriel [13] for $d \geqslant 2$, that the following conditions are equivalent:

(i) $E(\sup_{t \in Q} |X_t|) < \infty$;

(ii) $E(\sup_{t \in Q} t^{-1} |Y_t|) < \infty$;

(iii) $E(|Y_t| \log^+ |Y_t|)^d) < \infty$.

In the case $d = 1$, results of B. Davis [1971], and McCabe and Shepp [1970], show that also the following conditions are equivalent:

(iv) $E|X_\tau| < \infty$ for each $\tau \in \Sigma$ (or each $\tau \in \bar{\Sigma}$)

(v) $E|\tau^{-1} Y_\tau| < \infty$ for each $\tau \in \Sigma$ (or each $\tau \in \bar{\Sigma}$).

Cairoli and Gabriel have shown that (v) for general $d \geqslant 2$ is equivalent with (iii)$_1$, i.e., with $E(|Y_t| \log^+ |Y_t|) < \infty$. The question whether (iv) for $d \geqslant 2$ is equivalent to (iii)$_1$ was posed. The following theorem contains a positive answer in the case $d = 2$, because then the independence of the Y_t implies that all stopping rules are given by tactics.

THEOREM 4.1. *Let $(Y_t, t \in Q = \mathbb{N}^d)$ be independent identically distributed random variables, then for every d,* (iii)$_1$ *is equivalent to*

(iv)$_{\mathrm{T}}$ $E|X_\tau| < \infty$ *for all* $\tau \in T$.

In fact, if (iii)$_1$ *holds then there exists an* $M < \infty$ *with* $E|X_\tau 1_{\{\tau < \infty\}}| \leqslant M$ *for all* $\tau \in \bar{T}$. *If* $Q = \mathbb{N} \times \mathbb{N}$, *then* (iii)$_1$ *is equivalent with* (iv).

Proof. T contains all stopping times taking values in $Q_0 = \{t \in \mathbb{N}^d : t_2 = \cdots = t_d = 1\}$. Thus, by the known special case $d = 1$, (iv)$_{\mathrm{T}}$ implies (iii)$_1$. Put $M = E(\sup_{t \in Q_0} |X_t|)$. The equivalence of (iii)$_1$ with (i) for the case $d = 1$ yields $M < \infty$. By the linear embedding theorem there exists for every $\tau \in \bar{T}$ a $\sigma \in \bar{T}$ taking values only in $Q_0 \cup \{\infty\}$, such that $E|X_\tau 1_{\{\tau < \infty\}}| = E|X_\sigma 1_{\{\sigma < \infty\}}|$. ∎

(c) Wald's equation: Usually Wald's equation is only stated in the case $E\tau < \infty$. Robbins ans Samuel [19] showed that the equation holds also in the case $E\tau = \infty$, except when $EY_t = 0$. We generalize their result as follows:

THEOREM 4.2. *Let Q be a locally finite partially ordered set and $p \leqslant q$ for all $q \in Q$. Let $(Y_t,\ t \in Q)$ be independent indentically distributed integrable random variables, and let $X_t = \sum_{s \leqslant t} Y_s$. If $\mathscr{H}$ is a tactic with $P(\tau_{\mathscr{H}} < \infty) = 1$ and $EX_{\tau_{\mathscr{H}}}$ is well-defined (i.e., $EX^+_{\tau_{\mathscr{H}}} < \infty$ or $EX^-_{\tau_{\mathscr{H}}} < \infty$), then $EX_{\tau_{\mathscr{H}}} = E\,|\tau_{\mathscr{H}}| \cdot EY_p$ holds provided that either $EY_p \neq 0$ or $E\,|\tau_{\mathscr{H}}| < \infty$. In order that $EX_{\tau_{\mathscr{H}}}$ exists, it suffices that EY_p and $E\,|\tau_{\mathscr{H}}|$ both be finite. If $Q = \mathbb{N} \times \mathbb{N}$, then in this statement it is possible to replace $\tau_{\mathscr{H}}$ by σ, where σ is any stopping rule.*

Proof. Combine the linear embedding theorem with the special case $Q = \mathbb{N}$ due to Robbins and Samuel. ∎

In exactly the same way one obtains a generalization of Wald's identity. (See, e.g., Breiman [2, p. 100] for the case $Q = \mathbb{N}$.)

THEOREM 4.3. *Let Q and $(Y_t,\ t \in Q)$ be as in Theorem 4.2. Assume that for some real $\lambda \neq 0$, $\varphi(\lambda) = Ee^{\lambda Y}$ exists and $\varphi(\lambda) \geqslant 1$. Let $\mathscr{H}$ be a tactic such that $X_t = \sum_{s \leqslant t} Y_s$ satisfies $|X_t| < \gamma < \infty$ for $t \leqslant \tau_{\mathscr{H}}$, and such that $E\,|\tau_{\mathscr{H}}| < \infty$. Then*

$$E(e^{\lambda X_{\tau_{\mathscr{H}}}}/\varphi(\lambda)^{\tau_{\mathscr{H}}}) = 1.$$

Another immediate consequence of the linear embedding theorem is:

THEOREM 4.4. *Let Q, $Y_t(t \in Q)$ and X_t be as in Theorem 4.2. If $EY_t = 0$ and $E\,|\tau_{\mathscr{H}}| < \infty$, then $EX^2_{\tau_{\mathscr{H}}} = E\,|\tau_{\mathscr{H}}| \cdot EY^2_p$.*

Proof. Again combine the linear embedding theorem with the 1-dimensional case due to Chow, Robbins and Teicher [9] (see also Blackwell and Girshick [1] and Wolfowitz [25]).

It should be pointed out that Wald's equation in the usual form, i.e., assuming $E\tau < \infty$, holds for all $\tau \in \Sigma$.

5. CONSTRUCTION OF ORDERED STOPPING RULES FOR INDEPENDENT PROCESSES

Let Q be a countable partially ordred set which is locally finite. Recall that this means that for any $t \in Q$ there exist only finitely many $s \in Q$ with $s \leqslant t$.

Let $X = (X_t, t \in Q)$ be a family of independent real-valued integrable random variables X_t defined on a probability space $(\Omega, \mathfrak{F}, P)$. Let $\mathfrak{F}_t$ be the σ-algebra $\sigma(X_s, s \leqslant t)$ generated by all X_s with $s \leqslant t$.

Σ' denotes the set of *ordered* stopping rules τ': $\tau' \in \Sigma$ is called *ordered* if there exist a totally ordered subset $Q' \subset Q$ with $P(\tau' \in Q') = 1$. Call $\tau' \in \Sigma'$ an *ordered stopping rule in the strict sense* ($\tau' \in \Sigma''$) if in addition $\{\tau' = t\} \in \sigma(X_s, s \leqslant t, s \in Q')$ for all $t \in Q'$.

THEOREM 5.1. *Let $(X_t, t \in Q)$ be as above. Assume $E(\sup_{t \in Q} X_t^-) < \infty$. If $\tau \in \Sigma$ satisfies $EX_\tau < \infty$, then there exists a $\tau' \in \Sigma''$ with $EX_{\tau'} \geqslant EX_\tau$.*

Proof. We first assume that $\Omega = \mathbb{R}^Q$, with the product Borel-σ-algebra. P is the product of the distributions P_t of X_t, $t \in Q$ and X_t is the coordinate variable projecting $\omega \in \Omega$ onto its tth coordinate ω_t.

A measurable map $\tau: \Omega \to Q \cup \{\infty\}$ with $P(\tau = \infty) = 0$ is an element of Σ if the sets $\{\tau = t\} = B_t = B_t(\tau)$ belong to $\mathfrak{F}_t$. Conversely, if $\{B_t\}$ $(t \in Q \cup \{\infty\})$ is a disjoint measurable partition of Ω with $B_t \in \mathfrak{F}_t$ and $P(B_\infty) = 0$, then this defines a stopping time by $\{\tau = t\}$ on B_t.

Finite Q. We start with the case where Q is finite. After that we shall explain which additional arguments are used for the general case.

If $Q_0 \subset Q$, we shall write $\Omega_{Q_0} = \mathbb{R}^{Q_0}$.

P_{Q_0} denotes the distribution of the process $(X_t, t \in Q_0)$: $\Omega \to \Omega_{Q_0}$, i.e., the product of the distributions P_t of X_t $(t \in Q_0)$. For $B \subset \Omega$ and $\eta \in \Omega_{Q_0}$, write $B(\eta) = \{\rho \in \Omega_{Q_0^c}: \rho = (\rho_t, t \in Q_0^c), (\eta, \rho) \in B\}$. Here (η, ρ) denotes the element $\omega \in \Omega$ with $\omega_t = \eta_t$ $(t \in Q_0)$ and $\omega_t = \rho_t$ $(t \in Q_0^c)$.

If it is necessary for clarity, we may also write $(\eta, \rho)_{Q_0}$ for (η, ρ).

If $(B_t, t \in Q \cup \{\infty\})$ is a family defining a stopping time τ in Ω by $\{\tau = t\}$ on B_t, and $P(B_t) = 0$ for $t \in Q_0$, we can change the family on a null set by adding the sets B_t $(t \in Q_0)$ to B_∞. This does not change EX_τ. We can therefore change τ on a null set and assume $B_t = \varnothing$ for $t \in Q_0$.

For any $\eta \in \Omega_{Q_0}$, the sets $B_t(\eta)$ with $t \in Q_0^c$ belong to the σ-algebra generated by the coordinate variables in $Q_{Q_0^c}$ which are indexed by $s \leqslant t$, so that the family $B_t(\eta)$ defines a stopping time if $P_{Q_0^c}(B_\infty(\eta)) = 0$. This happens for almost all η and we shall always assume in our constructions that the exceptional null sets of η's are avoided. Denote the stopping time in $\Omega_{Q_0^c}$ constructed in this way by τ^η. τ^η can also be viewed as a stopping time in Ω (with the sets $\Omega_{Q_0} \times B_t(\eta)$) which does not depend on the coordinates X_t with $t \in Q_0$.

Now, let $Q_0 = \{t \in Q: P(\tau = t) = 0\}$. Then given τ the above construction yields stopping times τ^η which do not depend anymore on the variables X_t, $t \in Q_0$. We claim that $EX_\tau \leqslant EX_{\tau^\eta}$ for at least one $\eta \in \Omega_{Q_0}$: We have

$$EX_\tau = \int_{\Omega_{Q_0}} \int_{\Omega_{Q_0^c}} X_\tau(\eta, \rho) \, P_{Q_0^c}(d\rho) \, P_{Q_0}(d\eta).$$

The construction of τ^η is such that $(X_{\tau^\eta})(\eta', \rho) = X_\tau(\eta, \rho)$ for all $\eta' \in \Omega_{Q_0}$. Therefore there must exist at least one η with

$$\int_{\Omega_{Q_0^c}} X_{\tau^\eta}(\eta', \rho) P_{Q_0^c}(d\rho) \geqslant EX_\tau \qquad \text{for all } \eta'.$$

But this means $EX_{\tau^\eta} \geqslant EX_\tau$. (So far we have only eliminated some randomization.)

Let τ_1 denote the new stopping time.

Put $Q_1 = \{t \in Q: P(\tau_1 = t) = 0\}$. Then $Q_1 \supset Q_0$, because the sets $B_t(\eta)$ above were empty for $t \in Q_0$. We can now repeat the procedure to obtain a stopping time τ_2 not depending on the coordinates X_t $(t \in Q_1$. As Q is finite, we shall obtain, after finitely many applications of this procedure, a stopping time $\tau_{r_1} = \nu$ with

$$EX_\nu \geqslant EX_\tau \tag{5.1}$$

such that with $Q_{r_1} = \{t \in Q: P(\nu = t) = 0\}$ we have

$$\{\nu = t\} \in \sigma(X_s, s \leqslant t, s \in Q_{r_1}^c) \tag{5.2}$$

for all $t \in Q_{r_1}^c$. In fact we may assume that $\{\nu = t\} = \varnothing$ for $t \in Q_{r_1}$.

Second Step. There now exists a unique point $u \in Q_{r_1}^c$ such that u is a minimal point in $Q_{r_1}^c$ (not necessarily in Q), (minimal means $v \leqslant u$, $v \in Q_{r_1}^c \Rightarrow v = u$). Indeed, if there existed two minimal points $u_1, u_2, u_1 \neq u_2$, then we would have

$$P(\nu = u_1) > 0, \qquad P(\nu = u_2) > 0.$$

But $\{\nu = u_1\}$ would be measurable in $\sigma(X_{u_1})$, $\{\nu = u_2\}$ would be measurable in $\sigma(X_{u_2})$. The independence of $\sigma(X_{u_1})$ and $\sigma(X_{u_2})$ would imply $P(\{\nu = u_1\} \cap \{\nu = u_2\}) > 0$, a contradiction. The uniqueness of u also implies $v \geqslant u$ for all $v \in Q_{r_1}^c$.

Put $A_t = \{\nu = t\}$, $R_0 = \{u\}$. We can consider Ω as the product $\Omega_{R_0} \times \Omega_{R_0^c} = \{(\eta, \rho): \eta \in \Omega_{R_0}, \rho \in \Omega_{R_0^c}\}$; i.e., the coordinates are now split into two groups in a different way. A_u is of the form $F_u \times \Omega_{R_0^c}$ for some Borel set $F_u \subset \mathbb{R}^1$, because A_u depends only on X_u.

For $\eta \in F_u^c$, $v \in Q_{r_1}^c$ with $v > u$, the sets

$$A_v(\eta) = \{\rho \in \Omega_{R_0^c}: (\eta, \rho) \in A_v\}$$

can now be used to define a stopping rule ν^η corresponding to the family of sets $\{A_u, F_u^c \times A_v(\eta)(v > u)\}$. Any such ν^η has the following property: If

$X_u(\omega) \in F_u$, then $\nu^\eta(\omega) = u$; if $X_u(\omega) \notin F_u$, then $\nu^\eta(\omega) > u$ and the value of $\nu^\eta(\omega)$ then does not depend anymore on the value which $X_u(\omega)$ assumes in F_u^c.

We claim that for at least one ν^η, $EX_{\nu^\eta} \geqslant EX_\nu$: In fact

$$EX_\nu = \int_{\{\nu = u\}} X_\nu \, dP + \int_{F_u^c} \left\{ \int_{\Omega_{R_0^c}} X_\nu(\eta, \rho) \, P_{R_0^c}(d\rho) \right\} P_{R_0}(d\eta).$$

There must be at least one $\eta \in F_u^c$ such that

$$EX_\nu \leqslant \int_{\{\nu = u\}} X_\nu \, dP + P(X_u \in F_u^c) \left\{ \int_{\Omega_{R_0^c}} X_\nu(\eta, \rho) \, P_{R_0^c}(d\rho) \right\}.$$

By the construction of ν^η, we have $X_{\nu^\eta}(\eta', \rho) = X_\nu(\eta, \rho)$ for all $\eta' \in F_u^c$. Therefore the right-hand side equals EX_{ν^η}.

On $\{\nu = u\}$ we have $\nu^\eta = \nu$, and on this set we shall not make any further modifications so that the desired ordered stopping time τ' will also take the value u on $\{\nu = u\}$. The above inequalities imply

$$E(X_{\nu^\eta} \mid \nu^\eta > u) \geqslant E(X_\nu \mid \nu^\eta > u).$$

We now start over again the same construction on the set $\{\nu^\eta > u\}$, with a renormalized probability measure. The restriction of ν^η to this set will play the same role as τ played in the first step, so that we now first eliminate the influence on the stopping by those states $v > u$ at which ν^η stops with probability 0. Then there will again be one minimal state among the $v > u$ at which the modification of ν^η stops with positive probability. This will be the second state on which τ' will stop with positive probability, etc.

As Q is finite, the iteration of this procedure ends after finitely many steps. With each step we have only increased the expected value of the stopped process.

It remains to eliminate the condition $\Omega = \mathbb{R}^Q$ to finish the proof of the finite case; this will be done later.

Infinite Q. 1. *First Case*: τ is optimal. (As before $EX_\tau < \infty$ and $E \sup_{t \in Q} X_t^- < \infty$.) We again assume $\Omega = \mathbb{R}^Q$.

The basic idea of the construction is the same as in the finite case, but there are two new difficulties.

(1) One gets a sequence of modifications, but these may lead to stopping on ever and ever larger t's. Thus in the limit one may not be able to stop with probability 1.

(2) Even if one stops with probability 1 and each of the modifications is optimal, the limit may not be good.

If τ is optimal, this difficulty can be overcome as follows. In this case almost all choices of the η's will be equally good as far as EX_{τ^η} is concerned. We can therefore alternate two types of selection of the η's: In one sequence of steps we make sure that the ultimate stopping time stops on almost all of Ω; in between we put steps which make $EX_{\tau'}$ large.

For sufficiently large N_1 there exist $t_1,..., t_{N_1} \in Q$ with

$$P\left(\bigcup_{i=1}^{N_1} \{\tau = t_i\}\right) > 2^{-1}.$$

Put $J_1 = \{t_1,..., t_{N_1}\}$. As Q is locally finite, we may assume $s \in J_1$ for all s which are $\leqslant t_i$ for at least one t_i $(i = 1,..., N_1)$. Let $Q_0 = \{s \in J_1: P(\{\tau = s\}) = 0\}$. We use Q_0 as in the first step of the finite case. Again, let $B_t = \{\tau = t\}$. We claim that

$$\int_{\Omega_{Q_0^c}} X_\tau(\eta, \rho)\, P_{Q_0^c}(d\rho) = EX_\tau$$

for P_{Q_0}-almost all η: If there is an η for which the integral is $> EX_\tau$, then $EX_{\tau^\eta} > EX_\tau$. This contradicts the optimality of τ. Put $g_1(\eta) = \int_{\Omega_{Q_0^c}} 1_{\{\tau \in J_1\}}(\eta, \rho)\, P_{Q_0^c}(d\rho)$. We have $P(\tau^\eta \in J_1) = g_1(\eta)$. Because of $P(\tau \in J_1) = \int g_1(\eta)\, P_{Q_0}(d\eta)$, we can pick η in such a way that $P(\tau^\eta \in J_1) \geqslant P(\tau \in J_1)$.

We now go through the same constructions as in the finite case with $Q_1 = \{t \in J_1: P(\tau_1 = t) = 0\}$, etc. The optimality of τ (and therefore of all modifications) insures that almost all "fibers" give the same contribution to the expected value of the stopped variables so that the "fibers" can be selected so as to make stopping on J_1 at least as likely as for the previous modification. When ν is constructed, we can be sure that the unique minimal point u belongs to J_1, because there is stopping in J_1. Now one can do the second step of the construction, and in finitely many steps one arrives at a stopping time μ (corresponding to the τ' of the finite case) which has the following properties:

(i) μ stops with probability $\geqslant 2^{-1}$ in J_1 and the values which μ takes in J_1 are totally ordered: $q^{(1)} < q^{(2)} < \cdots < q^{(m_1)}$.

(ii) $EX_\mu = EX_\tau$.

(iii) All other values which μ takes with positive probability are $> q^{(m_1)}$.

(iv) The restriction of μ to $\{\mu > q^{(m_1)}\}$ depends only on $\sigma(X_t, t > q^{(m_1)})$.

The ultimate τ' promised in the statement of the theorem will coincide with μ on $\{\mu \leqslant q^{(m_1)}\}$. Therefore we now only look at μ on $\{\mu > q^{(m_1)}\}$. As we want to do similar things as above with this restriction of μ, we simply call it τ, and we write Ω for $\{\mu > q^{(m_1)}\}$. This τ also is optimal.

For sufficiently large $N_2 > N$ we find $t_{N_1+1},\dots, t_{N_2} \in Q$ with $t_i > q^{(m_1)}$ $(i = N_1 + 1,\dots, N_2)$ such that for $J_2 = \{t_{N_1+1},\dots, t_{N_2}\}$

$$\int_{\{\tau \notin J_2\}} |X_\tau| < \varepsilon_1$$

holds for some small $\varepsilon_1 > 0$. We may assume $s \in J_2$ for all $s \in Q$ which satisfy $q^{(m_1)} < s < t_k$ for some k. We have

$$\int_{\{\tau \in J_2\}} X_\tau \geqslant EX_\tau - \varepsilon_1 .$$

Put $Q_0 = \{s > q^{(m_1)}: P(\tau = s) = 0\}$, etc., and define $g_2(\eta)$ by

$$g_2(\eta) = \int_{\Omega_{Q_0^c}} 1_{\{\tau \in J_2\}} X_\tau(\eta, \rho)\, P_{Q_0^c}(d\rho).$$

Then $\int_{\{\tau \in J_2\}} X_\tau = \int_{\Omega_{Q_0}} g_2(\eta)\, P_{Q_0}(d\eta)$. As above the optimality of τ implies that $EX_{\tau^\eta} = EX_\tau$ for all η. The form of τ^η implies that $\int_{\{\tau^\eta \in J_2\}} X_{\tau^\eta} = g_2(\eta)$. It is therefore possible to choose an η with

$$\int_{\{\tau^\eta \in J_2\}} X_{\tau^\eta} \geqslant EX_\tau - \varepsilon_1 \qquad \text{and} \qquad EX_{\tau^\eta} = EX_\tau .$$

Again we can go through the constructions given in the finite case, this time making sure that each of the stopping times σ obtained in the course of the construction should satisfy

$$\int_{\{\sigma \in J_2\}} X_\sigma \geqslant EX_\tau - \varepsilon_1 \qquad \text{and} \qquad EX_\sigma = EX_\tau .$$

When all the elements of J_2 have been taken care of (either eliminating the dependence of the stopping times on X_t if t was an index where stopping occurred with probability 0, or stopping in t if t was one of the "lowest points"), we have arrived at a μ such that

(i) $\int_{\{\mu \in J_2\}} X_\mu \geqslant EX_\tau - \varepsilon_1$ and the values which μ takes in J_2 are totally ordered: $q^{(m_1+1)} < \cdots < q^{(m_2)}$ (with $q^{(m_1+1)} > q^{(m_1)}$).

(ii) $EX_\mu = EX_\tau$.

This is now true for the restrictions, but if one puts it together with what had been achieved previously, it remains true for the original τ.

Finally the statements analogous to (iii) and (iv) above are true.

We have used the elements $t \in J_1$ to assume the stopping on a large part of the space, and those of J_2 to nail down much of EX_τ. We now repeat the construction used with J_1 to obtain stopping on at least $\frac{3}{4}$ of the probability space by looking at some J_3 and then nail down more of EX_τ, etc. In the limit we arrive at an ordered stopping time τ' with $EX_{\tau'} = EX_\tau$.

Second Case. τ is *not* optimal. Again we assume $\Omega = \mathbb{R}^Q$. There exists a $\sigma \in \Sigma$ and $\varepsilon > 0$ with $EX_\sigma > EX_\tau + 2\varepsilon$. Let $c = E(\sup X_t^-)$. Choose a $c_1 > c$ so large that $P(B) \leqslant c/c_1 \Rightarrow \int_B \sup X_t^- < \varepsilon$. Put $X_\infty = -c_1$. One can choose finitely many $t_1, t_2, \ldots, t_N \in Q$ such that

$$\sum_{i=1}^{N} E(X_{t_i} 1_{\{\sigma = t_i\}}) - c_1 P(\sigma \notin \{t_1, \ldots, t_N\}) > EX_\tau + \varepsilon.$$

Put

$$\begin{aligned} \sigma_0 &= t_i && \text{on } \{\sigma = t_i\}\ i = 1, \ldots, N \\ &= \infty && \text{elsewhere.} \end{aligned}$$

Then $EX_{\sigma_0} > EX_\tau + \varepsilon$. Given σ_0, find an ordered stopping time σ_0' with values in $\{t_n, \ldots, t_N, \infty\}$ by the application of the finite case. (Note that Q is locally finite; the fact that there are infinitely many $t \in Q$ with $t \leqslant \infty$ does not matter because σ_0 takes only finitely many values and the dependence of the stopping times on the remaining $t \notin \{t_1, \ldots, t_N, \infty\}$ is eliminated at the beginning.) Now $EX_{\sigma_0'} \geqslant EX_{\sigma_0}$ implies

$$EX_{\sigma_0'} \geqslant EX_\tau + \varepsilon.$$

Let w denote the largest finite value of the ordered stopping time σ_0'. Define an ordered stopping time, which does not take on the value ∞, by

$$\begin{aligned} v &= \sigma_0' && \text{on} \quad \{\sigma_0' < \infty\} \\ &= w && \text{on} \quad \{\sigma_0' = \infty\}. \end{aligned}$$

To prove $EX_v \geqslant EX_\tau$, we consider two cases:

(a) Assume $EX_{\sigma_0'}^+ \geqslant EX_\tau + c$. Then

$$EX_v = \int_{\{\sigma_0' < \infty\}} X_{\sigma_0'} + \int_{\{\sigma_0' = \infty\}} X_w \geqslant EX_{\sigma_0'}^+ - c \geqslant EX_\tau.$$

Here the second inequality holds because $EX_{\sigma_0'}^+ = \int_{\{\sigma_0' < \infty\}} X_{\sigma_0'}^+$ and $\int_{\{\sigma_0' < \infty\}} X_{\sigma_0'}^- - \int_{\{\sigma_0' = \infty\}} X_w \leqslant c$.

(b) Assume $EX^+_{\sigma_0'} \leqslant EX_\tau + c$. Now $EX_{\sigma_0'} \geqslant EX_\tau$ implies $EX^-_{\sigma_0'} \leqslant c$. Hence $c_1 P(\sigma_0' = \infty) \leqslant c$, which implies

$$P(\sigma_0' = \infty) \leqslant c/c_1.$$

We obtain

$$EX_v \geqslant EX_{\sigma_0'} + \int_{\{\sigma_0' = \infty\}} X_w \geqslant EX_\tau + \varepsilon - \int_{\{\sigma_0' = \infty\}} X_w^- \geqslant EX_\tau$$

(by the choice of c_1).

(The proof for the case where τ is not optimal can be considerably simplified if Q is filtering to the right.)

We now show how to reduce the general case of the theorem to the case $\Omega = \mathbb{R}^Q$. The following proposition applies.

PROPOSITION 5.2. *Let Q be a countable partially ordered locally finite set. Let $(\mathscr{G}_t,\ t \in Q)$ be independent σ-algebras, $\mathfrak{F}_t = \sigma(\mathscr{G}_s, s \leqslant t)$, and set $\mathfrak{F}_\infty = \sigma(\mathscr{G}_s,\ s \in Q)$. Let $(Y_t, t \in Q)$ be independent random variables, $\mathscr{G}_t = \sigma(Y_t)$. Consider the space $\tilde{\Omega} = \mathbb{R}^Q$ with coordinate variables X_t, and let $\tilde{P}$ on $\tilde{\Omega}$ be the product measure of distributions of Y_t. Let Y be the map $Y: \Omega \to \mathbb{R}^Q$ with $Y_t(\omega) = X_t(Y(\omega))$. Then there exists a stopping rule $\tilde{\tau}$ on $\tilde{\Omega}$ with respect to the $\tilde{\mathfrak{F}}_\tau$ generated by X_s, $s \leqslant t$, such that almost surely*

$$\tau(\omega) = \tilde{\tau}(Y(\omega)).$$

Proof. Put

$$\mathscr{B}^t = \text{Borel sets in } \mathbb{R}^{K(t)},$$

$$\tilde{\mathscr{B}}^t = \text{cylinder sets in } \mathbb{R}^Q \text{ with basis in } \mathscr{B}^t,$$

$$X^t = \{X_s, s \leqslant t\}: X^t: \mathbb{R}^Q \to \mathbb{R}^{K(t)}.$$

(Thus the sets in $\tilde{\mathscr{B}}^t$ have the form $\{X^t \in B^t\} =: \tilde{B}^t$ with $B^t \in \mathscr{B}^t$.) As each $\{\tau = t\}$ is measurable in $\mathfrak{F}_t$, there exists a set $C_t \in \mathscr{B}^t$ with

$$\begin{aligned} \{\tau = t\} &= \{Y^t \in C_t\} \\ &= \{Y \in \tilde{C}_t\}. \end{aligned}$$

$Y^t: \Omega \to \mathbb{R}^{K(t)}$ is the process $\{Y_s, s \leqslant t\}$.

Of course, we would like to define $\tilde{\tau}$ by putting $\tilde{\tau}(\tilde{\omega}) = t$ on $\tilde{C}_t$, but the sets $\tilde{C}_t$ may not be disjoint.

We shall now replace the sets $\tilde{C}_t$ by smaller sets, which—in the end of a sequence of modifications—will be disjoint. We shall call also the modifications $\tilde{C}_t$ to avoid new notation. The changes will be such that the new $\tilde{C}_t$'s will differ from the old ones only on sets of measure 0 under the distribution of Y. For the new sets we shall still have $\{\tau = t\} = \{Y \in \tilde{C}_t\}$ a.e.

For any $t \in Q$ with $P(\tau = t) = 0$, we can take $\tilde{C}_t = \varnothing$ for the modification. For any $V \subset Q$ let P_V denote the distribution of $\{Y_t, t \in V\}$ on $\mathbb{R}^V$. For any fixed $u \neq v$ with $P(\tau = u) > 0$, $P(\tau = v) > 0$ we consider $V = V_{u,v} = K(u) \cap K(v)$, $V' = K(u) \backslash V$, $V'' = K(v) \backslash V$. If $V = \varnothing$, $\{\tau = u\}$ is independent of $\{\tau = v\}$, but $\{\tau = v\} \cap \{\tau = u\}$ is empty so that this cannot happen. For $\eta \in \mathbb{R}^V$ put

$$C_{u,\eta} = \{\rho \in \mathbb{R}^{V'}: (\eta, \rho)_V \in C_t\}.$$

Let $N_u = \{\eta \in \mathbb{R}^V: P_{V'}(C_{u,\eta}) = 0\}$. Omit all points (η, ρ) from C_u with $\eta \in N_u$. This is only a modification by a null set. Carry out the same reduction for C_v. Now any $\eta \in \mathbb{R}^V$ has the property that $C_{u,\eta}$ is empty or has positive $P_{V'}$-measure and lso $C_{v,\eta}$ is empty or has positive $P_{V''}$-measure.

For any $\eta \in \mathbb{R}^V$, $\tilde{C}_u \cap \tilde{C}_v$ contains all points $\tilde{\omega} \in \mathbb{R}^Q$ with $X_V(\tilde{\omega}) = \eta$, $X_{V'}(\tilde{\omega}) \in C_{u,\eta}$ and $X_{V''}(\tilde{\omega}) \in C_{v,\eta}$. Thus, if $\{\eta \in \mathbb{R}^V: C_{u,\eta} \neq \varnothing$ and $C_{v,\eta} \neq \varnothing\} = M_{uv}$ has positive P_V-measure, then $\tilde{C}_u \cap \tilde{C}_v$ must have positive P_Q-measure. This is impossible because

$$P_Q(\tilde{C}_u \cap \tilde{C}_v) = P(Y \in \tilde{C}_u \cap \tilde{C}_v) = P(\{\tau = u\} \cap \{\tau = v\}) = 0.$$

Therefore, if we omit all $(\eta, \rho) \in \mathbb{R}^V \times \mathbb{R}^{V'}$ with $\eta \in M_{uv}$ from C_u and all $(\eta, \rho) \in \mathbb{R}^V \times \mathbb{R}^{V''}$ with $\eta \in M_{u,v}$ from C_v, this is only a modification by a null set. But after this modification $\tilde{C}_u$ and $\tilde{C}_v$ are disjoint.

If we apply the same modification for all pairs (u, v) with $u \neq v$, we reduce the sets $\tilde{C}_t$ $(t \in Q)$ to disjoint sets while preserving the measurability in the σ-algebra $\tilde{\mathfrak{F}}_t$. We can now put $\tilde{\tau}(\tilde{\omega}) = t$ on $\tilde{C}_t$. The stopping time $\tau'(\omega) = \tilde{\tau}(Y(\omega))$ agrees with τ a.e.

Example. There exists a partially ordered locally finite countable set Q, independent random variables $X_t \geqslant 0$ $(t \in Q)$ and a stopping time τ with $EX_\tau = \infty$ such that for no ordered stopping time τ' one has $EX_{\tau'} = \infty$.

Proof. Let $Q = \mathbb{N} \times \{0, 1\}$ with the partial ordering given by $(k, 0) \leqslant (l, 0)$ if $k \leqslant l$, $(k, 1) \geqslant (k, 0)$. All $(k, 1)$ will be incomparable.

The random variables X_t, $t \in Q$, will be independent, and we assume $P(X_{k,0} = 0) = P(X_{k,0} = 1) = \frac{1}{2}$. The random variables $X_{k,1}$ will take the constant values $c_k \geqslant 0$. If c_k converges to ∞, then there exists a τ with

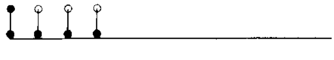

FIGURE 3

$EX_\tau = \infty$. Yet there cannot exist an ordered stopping time τ' with $EX_{\tau'} = \infty$. Indeed, any such stopping time would have to take at least 2 of the values $(k, 1)$ with positive probability (in fact countably many), and these are incomparable.

6. Dominated Estimates in Terms of Stopping Rules

These inequalities, proved in the linear case in Krengel and Sucheston [15], are of the form

$$\text{(M)} \qquad E(\sup_{t \in Q} X_t) \leqslant KV,$$

where $V = \sup_{\tau \in \Sigma} EX_\tau$ is the value of the process, and K is a universal constant. If the random variables X_t are measurable with respect to independent σ-algebras $\mathscr{G}_t$, and the stopping rules are with respect to $\mathfrak{F}_t = \sigma(\mathscr{G}_s,\ s \leqslant t)$, then the constant K is 2; if X_t are *averages* of positive random variables measurable with respect to $\mathscr{G}_s$, the constant K is $\leqslant 5.5$.

We show that no such constants exist in the case $Q = \mathbb{N}^2$. However, if the stopping rules in the usual sense are replaced by a wider class, called by Cairoli and Gabriel [5] "wide sense stopping points," and by Walsh [24] "weak stopping points," then the linear inequalities extend with the same constants. We will use the term *wide sense stopping rules* (or *times*). We at first show that (M) fails for ordinary stopping rules.

In fact, in the two-dimensional case (M) does not hold even when the X_t are indicator functions of sets A_t. (In the one-dimensional case one can easily obtain the constant 1 in that case, by stopping as soon as an X_t with value 1 is observed.) To see this, let A_t be independent sets with $P(A_t) = \alpha > 0$, $t \in \{1, 2, \dots, n\}^2$, and $P(A_t) = 0$ elsewhere, where α is very small. For α close to 0, $E \sup X_t \approx n^2\alpha$ because the independent sets A_t are nearly disjoint. On the other hand, $EX_\tau \leqslant 2n\alpha$ holds for all $\tau \in \Sigma$, because by Theorem 5.1 for any τ there is an ordered stopping time τ' with $EX_\tau \leqslant EX_{\tau'}$ and any ordered stopping time τ' can have $P(\tau' = t) > 0$ for at most $2n$ values of $t \in \{1, 2, \dots, n\}^2$.

For $t = (t_1, \dots, t_d) \in \mathbb{N}^d$ put $|t| = t_1 \cdot t_2 \cdot \cdots \cdot t_d$. Using known results on optimal stopping of $Y_t/|t|$ it is now also easy to give an example of independent random variables $0 \leqslant X_t$ $(t \in \mathbb{N}^2)$ with $E \sup X_t = \infty$ such that $V < \infty$.

Let $Y \geqslant 0$ be a random variable with $E(Y \log^+ Y) < \infty$ and $E(Y(\log^+ Y)^2) = \infty$. Let Y_t, $t \in \mathbb{N}^2$, be independent copies of Y and put $X_t = Y_t/|t|$. By a result of Gabriel [13], $E \sup X_t = \infty$. On the other hand by our theorem on ordered stopping times (Theorem 5.1), or our theorem on the replacement of tactics by linear stopping times (Theorem 4.1), one has

$V \leqslant E \sup_{t \in \mathbb{N}} t^{-1} Y_{t,1}$, and this is finite by a theorem of Burkholder [4], since $EY \log^+ Y < \infty$.

In the same way one shows that in the $\mathbb{N} \times \mathbb{N}$ case no estimate (M) holds for processes of the form $X_t = 1/|t| \sum_{s \leqslant t} Y_s$ with independent $\mathscr{G}_s$-measurable r.v.'s, Y_s.

We now discuss wide sense stopping rules. For $s = (s_1, ..., s_d)$, $t = (t_1, ..., t_d)$, $s \leqslant t$ means: $s_i \leqslant t_i$ for all i. $t + 1$ denotes $(t_1 + 1, t_2 + 1, ..., t_d + 1)$. Let $(\mathscr{G}_t, t \in Q)$ be independent σ-algebras. Let

$$\mathfrak{F}_t = \sigma(\mathscr{G}_s, s \leqslant t),$$
$$\mathfrak{F}_t^* = \sigma(\mathscr{G}_s, t + 1 \nleqslant s),$$
$$\mathfrak{F}_t^{**} = \sigma(\mathscr{G}_s, s_1 \leqslant t_1).$$

A random variable $\tau: \Omega \to Q \cup \{\infty\}$ almost surely taking values in Q is said to belong to Σ if for all t in Q, $\{\tau = t\} \in \mathfrak{F}_t$; to Σ^* if $\{\tau = t\} \in \mathfrak{F}_t^*$; to Σ^{**} if $\{\tau = t\} \in \mathfrak{F}_t^{**}$. (One could equivalently require that the sets $\{\tau \leqslant t\}$ belong to these σ-algebras.) The elements of Σ^* are called *wide sense stopping rules* (*times*). We have

THEOREM 6.1. *Let $(X_t, t \in Q = N^d)$ be idependent random variables, X_t generating $\mathscr{G}_t$. Then*

$$E(\sup_{t \in Q} X_t) \leqslant 2V^{**} \leqslant 2V^*,$$

where $V^{**} = \sup\{EX_\tau, \tau \in \Sigma^{**}\}$, and $V^* = \sup\{EX_\tau, \tau \in \Sigma^*\}$.

Proof. Let $Z_n = \sup\{X_t: t_1 = n\}$ and $\mathfrak{F}_n^Z = \sigma(\mathscr{G}_t: t_1 \leqslant n)$, then Z_n is independent of $\mathfrak{F}_{n-1}^Z$. By the result of Krengel and Sucheston [15], there exist stopping times ν (with respect to $\mathfrak{F}_n^Z$) such that $E \sup E_n \leqslant 2 \sup_\nu EZ_\nu$.

If $(\Omega', \mathfrak{F}', P')$ is a measure space, Q' countable and X'_s $(s \in Q')$ are $\mathfrak{F}'$-measurable integrable real-valued random variables, then there exist $\mathfrak{F}'$-measurable maps $\mu': \Omega' \to Q'$ such that $E \sup X'_s = EX'_\mu$.

Defining μ on Ω so that on $\{\nu = n\}$ it agrees with appropriate μ', we find that there exist maps $\nu: \Omega \to \mathbb{N}$ and $\mu: \Omega \to \mathbb{N}^{d-1}$ such that $\{\nu = n\} \in \mathfrak{F}_n^Z$, the restriction of μ to $\{\nu = n\}$ is $\mathfrak{F}_n^Z$-measurable, and

$$E \sup_{t \in Q} X_t = E \sup Z_n \leqslant 2 \sup_{(\nu,\mu)} EX_{(\nu,\mu)}.$$

The pairs (ν, μ) are stopping times belonging to Σ^{**}. ∎

The case of averages $X_t = (1/|t|) \sum_{s<t} Y_s$ with nonnegative independent random variables Y_s generating $\mathscr{G}_s$ is more interesting, because the Z_n need no longer be averages of independent random variables. However, in [3] the

inequality on averages has been extended to a class of processes closed under the formation of suprema, and this generalization enables us to treat averages in several dimensions:

Let $\mathcal{A}_i$, $i \in \mathbb{N}$, be independent σ-algebras and $\mathcal{A}(i,j) = \sigma(\mathcal{A}_k, i \leqslant k \leqslant j)$. A process Z_n $(n \in \mathbb{N})$ is called a process with independent *subadditive* nonnegative components if there exist random variables $Z_n(i,j) \geqslant 0$ ($n \in \mathbb{N}$, $1 \leqslant i \leqslant j < \infty$), measurable in $\mathcal{A}(i,j)$, such that $Z_n = Z_n(1,n)$, and the inequalities

$$Z_n(i,k) \leqslant Z_n(i,j) + Z_n(j+1,k) \qquad (i \leqslant j < k)$$

and

$$Z_n(i,j) \leqslant Z_n(i^*,j^*) \qquad (i^* \leqslant i \leqslant j \leqslant j^*)$$

hold. (This notion is due to Brunel and Krengel [3], but due to a misprint $\mathcal{A}(i,j)$ is defined by $\sigma(\mathcal{A}_k, i < k < j)$ in [3]). We have

THEOREM 6.2. *Let $\mathcal{G}_s$ $(s \in Q = \mathbb{N}^d)$ be independent σ-algebras, $\mathcal{A}_i = \sigma(\mathcal{G}_s\colon s_1 = i)$ and let $X_t(t \in Q)$ be a process such that for each $(t_2,\ldots,t_d) \in \mathbb{N}^{d-1}$, $X_{(n,t_2,\ldots,t_d)}$ is a process with independent subadditive nonnegative components. Then*

$$E \sup_{t \in Q} X_t \leqslant 2^{-1}(1+\sqrt{3})^{-1} V^{**} \leqslant 2^{-1}(1+\sqrt{3})^{-1} V^*.$$

Proof. We can apply the result of [3] to Z_u, to find good stopping times ν for the increasing family $\mathcal{A}(1,n)$ and Z_n. As all $X_{(n,t_2,\ldots,t_d)}$ are measurable with respect to $\mathcal{A}(1,n)$ we can then find maps $\mu\colon \Omega \to \mathbb{N}^{d-1}$ as above, such that μ is $\mathcal{A}(1,n)$-measurable on $\{\nu = n\}$, and such that $\tau = (\nu,\mu) \in \Sigma^{**}$ gives sufficiently large values to EX_τ.

EXAMPLE. The above inequality holds for processes of the form $X_t = ((1/|t|)\sum_{s \leqslant t} Y_s)^{1/2}$ and independent $Y_s \geqslant 0$.

Remarks Added August 1980

1. We mention in the introduction the notion of control variable introduced by Haggstrom [14] in the case of a tree. The definition, extended here to a general locally finite set Q, is as follows: A mapping $\tau\colon \Omega \to Q$ is a *control variable* if for all $s \in Q$ and all $t \in D_s$, one has $\{\tau \geqslant t\} \in \mathfrak{F}_s$. It is easy to see that every control variable is a stopping time, and furthermore every control variable τ defines a tactic $\mathscr{X}$ with $\tau = \tau_{\mathscr{X}}$. However the following example shows that in the case of $\mathbb{N} \times \mathbb{N}$ with σ-algebras generated by independent random variables, there is a tactic $\mathscr{X}$ such that the stopping time $\tau_{\mathscr{X}}$ is not a control variable. Thus the notion of a tactic is more general, and apparently more useful, than that of a control variable. Let $Q = \{0,1\}^2$, and let $\mathfrak{F}_t$, $t \in Q$, be generated by i.i.d. random variables Y_s, $s \leqslant t$, with $P(Y_s = 0) = P(Y_s = 1) = \frac{1}{2}$. Define a tactic $\mathscr{X}$ as follows: If $Y_{(0,0)} = 0$, go to $(1,0)$; if

$Y_{(0,0)} = 1$, go to $(0, 1)$. From $(0, 1)$ go deterministically to $(1, 1)$. From $(1, 0)$ go to $(1, 1)$ if $Y_{(1,0)} = 1$; to $(1, 0)$—that is, stop—if $Y_{(1,0)} = 0$. Then if $\tau = \tau_{\mathscr{H}}$ is the stopping time defined by $\mathscr{H}$, the set

$$\{\tau \geqslant (0, 1)\} = \{Y_{(0,0)} = 1\} \cup \{Y_{(0,0)} = 0,\ Y_{(1,0)} = 1\}$$

is not $\mathfrak{F}_{(0,0)}$ measurable, hence τ is not a control variable.

2. D. L. Burkholder and R. F. Gundy (*Acta Math.* **124**, 249–304, in particular p. 281) showed that if (Y_n) is a martingale difference sequence, $EY_k = 0$, $E(Y_k^2 \mid Y_1, Y_2, \ldots, Y_{k-1}) = 1$, and $\tau \in \Sigma$, $E\tau^{1/2} < \infty$, then $E(\sum_{k=1}^{\tau} Y_k) = 0$. In the particular case when the Y_n's are i.i.d. random variables, this result extends to tactics by the linear embedding theorem.

3. Professor A. Dvoretzky has pointed out to us that the 3 conditions (i), (ii), (iii) in the beginning of Section 2 are equivalent with

(1) For each j, $s^j \subset E_j \subset L(s_j)$,

(2) $M(s) \backslash \bigcap_j L(s^j) \subset \bigcup E_j$.

The proof of the equivalence of the two sets of conditions in simple.

References

[1] Blackwell, D. and Girshick, M. A. (1947). A lower bound for the variance of some unbiased sequential estimates. *Ann. Math. Statist.* **18** 215–230.

[2] Breiman, L. (1968). *Probability*. Addison-Wesley, Reading, Mass.

[3] Brunel, A. and Krengel, U. (1979). Parier avec un prophète dans le cas d'un processus sous-additif. *C. R. Acad. Sci. Paris Sér. A* **288** 57–60.

[4] Burkholder, D. L. (1962). Successive conditional expectations of an integrable function. *Ann. Math. Statist.* **33** 887–893.

[5] Cairoli, R. and Gabriel, J. P. (1977-78). Arrêt de certaines suites multiples de variables aléatoires independentes. In *Séminaire de Probabilités*. Lecture Notes in Mathematics Vol. 721, pp. 174–198. Springer-Verlag, New York/Berlin.

[6] Chow, Y. S. and Dvoretzky, A. (1969). Stopping rules for x_n/n and related problems. *Israel J. Math.* **7** 240–248.

[7] Chow, Y. S. and Robbins, H. E. (1965). On optimal stopping rules for S_n/n. *Illinois J. Math.* **9** 444–454.

[8] Chow, Y. S., Robbins, H. and Siegmund, D. (1971). *Great Expectations: The Theory of Optimal Stopping*. Houston Mifflin, Boston.

[9] Chow, Y. S., Robbins, H., and Teicher, H. (1965). Moments of randomly stopped sums. *Ann. Math. Statist.* **36** 789–799.

[10] Davis, B. (1971). Stopping rules for S_n/n and the class LlogL. *Z. Wahrsch. Verw. Gebiete* **17** 147–150.

[11] Davis, B. (1973). Moments of random walk having infinite variance and the existence of certain optimal stopping rules for S_n/n. *Illinois J. Math.* **17** 75–81.

[12] Dvoretzky, A. (1967). Existence and properties of certain optimal stopping rules. In *Proc. Berkeley Symposium Math. Statist. Prob.*, University of California Press, Berkeley and Los Angeles, pp. 441–452.

[13] Gabriel, J. P. (1977). Martingales with a countable filtering index set. *Ann. Prob.* **5** 888–898.

[14] HAGGSTROM, G. W. (1966). Optimal stopping and experimental design. *Ann. Math. Stat.* **37** 7–29.

[15] KRENGEL, U. AND SUCHESTON, L. (1978). On semiamarts, amarts, and processes with finite value. *Advances in Probab.* **4** 197–266.

[16] MCCABE, B. J. AND SHEPP, L. A. (1970). On the supremum of S_n/n. *Ann. Math. Statist.* **41** 2166–2168.

[17] MILLET, A. AND SUCHESTON, L. (1980). Convergence of classes amarts indexed by directed sets. *Canadian J. Math.* **32** 86–125.

[18] NEVEU, J. (1975). *Discrete Parameter Martingales.* North-Holland, Amsterdam.

[19] ROBBINS, H. E. AND SAMUEL, E. (1966). An extension of a lemma of Wald. *J. Appl. Probab.* **3** 272–273.

[20] SHIRYAEV, A. N. (1978). *Optimal Stopping Rules.* Springer-Verlag, New York, Heidelberg, Berlin.

[21] SMYTHE, R. T. (1974). Sums of independent random variables on partially ordered sets. *Ann. Probab.* **2** 906–917.

[22] SNELL, J. L. (1952). Application of martingale system theorems. *Trans. Amer. Math. Soc.* **73** 293–312.

[23] WALD, A. (1947). *Sequential Analysis.* Wiley, New York.

[24] WALSH, J. B. (1979). Convergence and regularity of multiparameter strong martingales, *Z. Wahrsch. Verw. Gebiete* **46** 177–192.

[25] WOLFOWITZ, J. (1947). The efficiency of sequential estimates and Wald's equation for sequential processes. *Ann. Math. Statist.* **18** 215–230.